Measuring Performance of Continuous-Time Stochastic Processes using Timed Automata

Tomáš Brázdil   Jan Krčál   Jan Křetínský

Antonín Kučera   Vojtěch Řehák

Faculty of Informatics, Masaryk University,
Botanická 68a, 60200 Brno,
Czech Republic
{brazdil, krcal, jan.kretinsky, kucera, rehak}
@fi.muni.cz

Abstract

We propose deterministic timed automata (DTA) as a model-independent language for specifying performance and dependability measures over continuous-time stochastic processes. Technically, these measures are defined as limit frequencies of locations (control states) of a DTA that observes computations of a given stochastic process. Then, we study the properties of DTA measures over semi-Markov processes in greater detail. We show that DTA measures over semi-Markov processes are well-defined with probability one, and there are only finitely many values that can be assumed by these measures with positive probability. We also give an algorithm which approximates these values and the associated probabilities up to an arbitrarily small given precision. Thus, we obtain a general and effective framework for analysing DTA measures over semi-Markov processes.

1   Introduction

Continuous-time stochastic processes, such as continuous-time Markov chains, semi-Markov processes, or generalized semi-Markov processes [23, 6, 20, 17], have been widely used in practice to determine performance and dependability characteristics of real-world systems. The desired behaviour of such systems is specified by various measures such as mean response time, throughput, expected frequency of errors, etc. These measures are often formulated just semi-formally and chosen specifically for the system under study in a somewhat ad hoc manner. One example of a rigorous and model-independent specification language for performance and dependability properties is Continuous Stochastic Logic (CSL) [3, 5] which allows to specify both steady state and transient measures over the underlying stochastic process. The syntax and semantics of CSL is inspired by

the well-known non-probabilistic logic CTL [13]. The syntax of CSL defines *state* and *path* formulae, interpreted over the states and runs of a given stochastic process $\mathcal{M}$. In particular, there are two probabilistic operators, $\mathcal{P}_{\bowtie \varrho}(\cdot)$ and $\mathcal{S}_{\bowtie \varrho}(\cdot)$, which refer to the transient and steady state behaviour of $\mathcal{M}$, respectively. Here $\bowtie$ is a numerical comparison (such as $\leq$) and $\varrho \in [0, 1]$ is a rational constant. If φ is a path formula[1] (which is either valid or invalid for every run of $\mathcal{M}$), then $\mathcal{P}_{\geq 0.7}(\varphi)$ is a state formula which says "the probability of all runs satisfying φ is at least 0.7". If Φ is a state formula, i.e., Φ is either valid or invalid in every state, then $\mathcal{S}_{\geq 0.5}(\Phi)$ is also a state formula which says "the π-weighted sum over all states where Φ holds is at least 0.5". Here π is the steady-state distribution of $\mathcal{M}$. The logic CSL can express quite complicated properties and the corresponding model-checking problem over continuous-time Markov chains is decidable. However, there are also several disadvantages.

(a) The semantics of steady state probabilistic operator $\mathcal{S}_{\bowtie \varrho}(\cdot)$ assumes the existence of invariant distribution which is not guaranteed to exist for all types of stochastic processes with continuous time (the existing works mainly consider CSL as a specification language for ergodic continuous-time Markov chains).

(b) In CSL formulae, all measures are explicitly quantified, and the model-checking algorithm just verifies constraints over these measures. Alternatively, we might wish to *compute* certain measures up to a given precision.

In this paper, we propose *deterministic timed automata (DTA)* [2] as a model-independent specification language for performance and dependability measures of continuous-time stochastic processes. The "language" of DTA can be interpreted over *arbitrary* stochastic processes that generate timed words, and their expressive power appears sufficiently rich to capture many interesting run-time properties (although we do not relate the expressiveness of CSL and DTA formally, they are surely incomparable because of different "nature" of the two formalisms). Roughly speaking, a DTA $\mathcal{A}$ "observes" runs of a given stochastic process $\mathcal{M}$ and "remembers" certain information in its control states (which are called *locations*). Since $\mathcal{A}$ is deterministic, for every run σ of $\mathcal{M}$ there is a unique computation $\mathcal{A}(\sigma)$ of $\mathcal{A}$, which determines a unique tuple of "frequencies" of visits to the individual locations of $\mathcal{A}$ along σ. These frequencies are the values of "performance measures" defined by $\mathcal{A}$ (in fact, we consider *discrete* and *timed* frequencies which are based on the same concept but defined somewhat differently).

Let us explain the idea in more detail. Consider some stochastic process $\mathcal{M}$ whose computations (or runs) are infinite sequences of the form $\sigma = s_0 \, t_0 \, s_1 \, t_1 \cdots$ where all s_i are "states" and t_i is the time spent by performing the transition from s_i to s_{i+1}. Also assume a suitable probability space defined over the runs of $\mathcal{M}$. Let Σ by a finite alphabet and L a labelling which assigns a unique letter $L(s) \in \Sigma$ to every state s of $\mathcal{M}$. Intuitively, the letters of Σ correspond to collections of

---

[1]In CSL, φ can be of the form $\mathcal{X}_I \Phi$ or $\Phi_1 \mathcal{U}_I \Phi_2$ where Φ, Φ_1, Φ_2 are state formulae, and $\mathcal{X}_I, \mathcal{U}_I$ are the modal connectives of CTL parametrized by an interval I. Boolean connectives can be used to combine just state formulae.

predicates that are valid in a given state. Thus, every run $\sigma = s_0 \, t_0 \, s_1 \, t_1 \cdots$ of $\mathcal{M}$ determines a unique *timed word* $w_\sigma = L(s_0) \, t_0 \, L(s_1) \, t_1 \cdots$ over Σ.

A DTA over Σ is a finite-state automaton $\mathcal{A}$ equipped with finitely many internal clocks. Each control state (or location) q of $\mathcal{A}$ has finitely many outgoing edges $q \longrightarrow q'$ labeled by triples (a, g, X), where $a \in \Sigma$, g is a "guard" (a constraint on the current clock values), and X is a subset of clocks that are reset to zero after performing the edge. A configuration of $\mathcal{A}$ is a pair (q, ν), where q and ν are the current location and the current clock valuation, respectively. Every timed word $w = c_0 \, c_1 \, c_2 \, c_3 \cdots$ over Σ (where $c_i \in \Sigma$ iff i is even) then determines a unique run $\mathcal{A}(w) = (q_0, \nu_0) \, (q_1, \nu_1) \, (q_2, \nu_2) \cdots$ of $\mathcal{A}$ where q_0 is an *initial* location, ν_0 assigns zero to every clock, and (q_{i+1}, ν_{i+1}) is obtained from (q_i, ν_i) either by performing the only enabled edge $q_i \longrightarrow q_{i+1}$ labeled by (c_i, g, X) if i is even, or by simultaneously increasing all clocks by c_i if i is odd.

As a simple example, consider the following DTA $\hat{\mathcal{A}}$ over the alphabet $\{a\}$ with one clock x and the initial location q_0:
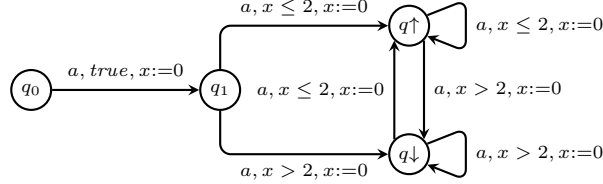
Intuitively, $\hat{\mathcal{A}}$ observes time stamps in a given timed word and enters either $q\uparrow$ or $q\downarrow$ depending on whether a given stamp is bounded by 2 or not, respectively. For example, a word $w = a \; 0.2 \; a \; 2.4 \; a \; 2.1 \cdots$ determines the run $\hat{\mathcal{A}}(w) = (q_0, 0) \, (q_1, 0) \, (q_1, 0.2) \, (q\uparrow, 0) \, (q\uparrow, 2.4) \, (q\downarrow, 0) \, (q\downarrow, 2.1) \cdots$

Let $w = a_0 \, t_0 \, a_1 \, t_1 \cdots$ be a timed word over Σ and q a location of $\mathcal{A}$. For every $i \in \mathbb{N}_0$, let $T^i(w)$ be the stamp t_i of w, and $Q^i(w)$ the location of $\mathcal{A}$ entered after reading the finite prefix $a_0 \, t_0 \cdots a_i$ of w. Further, let $1_q^i(w)$ be either 1 or 0 depending on whether $Q^i(w) = q$ or not, respectively. We define the *discrete* and *timed* frequency of visits to q along $\mathcal{A}(w)$, denoted by $\mathbf{d}_q^{\mathcal{A}}(w)$ and $\mathbf{c}_q^{\mathcal{A}}(w)$, in the following way (the '$\mathcal{A}$' index is omitted when it is clear from the context):

$$
\begin{aligned}
\mathbf{d}_q^{\mathcal{A}}(w) &= \limsup_{n \to \infty} \frac{\sum_{i=1}^{n} 1_q^i(w)}{n} \\
\mathbf{c}_q^{\mathcal{A}}(w) &= \limsup_{n \to \infty} \frac{\sum_{i=1}^{n} T^i(w) \cdot 1_q^i(w)}{\sum_{i=1}^{n} T^i(w)}
\end{aligned}
$$

Thus, every timed word w determines the tuple $\mathbf{d}^{\mathcal{A}}(w) = \left(\mathbf{d}_q^{\mathcal{A}}(w) \right)_{q \in Q}$ and the tuple $\mathbf{c}^{\mathcal{A}}(w) = \left(\mathbf{c}_q^{\mathcal{A}}(w) \right)_{q \in Q}$ of *discrete* and *timed $\mathcal{A}$-measures*, respectively.

DTA measures can encode various performance and dependability properties of stochastic systems with continuous time. For example, consider again the DTA $\hat{\mathcal{A}}$ above and assume that all states of a given stochastic process $\mathcal{M}$ are labeled with a. Then, the fraction

$$
\frac{\mathbf{d}_{q\uparrow}(w_\sigma)}{\mathbf{d}_{q\uparrow}(w_\sigma) + \mathbf{d}_{q\downarrow}(w_\sigma)}
$$

corresponds to the percentage of transitions of $\mathcal{M}$ that are performed within 2 seconds along a run σ. If $\mathcal{M}$ is an ergodic continuous-time Markov chain, then the above fraction takes the *same* value for almost all runs σ of $\mathcal{M}$. However, it makes sense to consider this fraction also for non-ergodic processes. For example, we may be interested in the expected value of $\mathbf{d}_{q\uparrow}/(\mathbf{d}_{q\uparrow} + \mathbf{d}_{q\downarrow})$, or in the probability of all runs σ such that the fraction is at least 0.5.

One general trouble with DTA measures is that $\mathbf{d}_q^{\mathcal{A}}(w)$ and $\mathbf{c}_q^{\mathcal{A}}(w)$ faithfully capture the frequency of visits to q along w only if the limits

$$\lim_{n\to\infty} \frac{\sum_{i=1}^n 1_q^i(w)}{n} \quad \text{and} \quad \lim_{n\to\infty} \frac{\sum_{i=1}^n T^i(w) \cdot 1_q^i(w)}{\sum_{i=1}^n T^i(w)}$$

exist, in which case we say that $\mathbf{d}^{\mathcal{A}}$ and $\mathbf{c}^{\mathcal{A}}$ are *well-defined* for w, respectively. So, one general question that should be answered when analyzing the properties of DTA measures over a particular class of stochastic processes is whether $\mathbf{d}^{\mathcal{A}}$ and $\mathbf{c}^{\mathcal{A}}$ are well-defined for almost all runs. If the answer is negative, we might either try to re-design our DTA or accept the fact that the limit frequency of the considered event simply does not exist (and stick to $\lim\sup$).

In this paper, we study DTA measures over semi-Markov processes (SMPs). An SMP is essentially a discrete-time Markov chain where each transition is assigned (apart of its discrete probability) a *delay density*, which defines the distribution of time needed to perform the transition. A computation (run) of an SMP $\mathcal{M}$ is initiated in some state s_0, which is also chosen randomly according to a fixed initial distribution over the state space of $\mathcal{M}$. The next transition is selected according to the fixed transition probabilities, and the selected transition takes time chosen randomly according to the density associated to the transition. Hence, each run of $\mathcal{M}$ is an infinite sequence $s_0\,t_0\,s_1\,t_1\cdots$, where all s_i are states of $\mathcal{M}$ and t_i are time stamps. The probability of (certain) subsets of runs in $\mathcal{M}$ is measured in the standard way (see Section 2).

The main contribution of this paper are general results about DTA measures over semi-Markov processes, which are valid for all SMPs where the employed density functions are bounded from zero on every closed subinterval (see Section 2). Under this assumption, we prove that for every SMP $\mathcal{M}$ and every DTA $\mathcal{A}$ we have the following:

(1) Both discrete and timed $\mathcal{A}$-measures are well defined for almost all runs of $\mathcal{M}$.

(2) Almost all runs of $\mathcal{M}$ can be divided into finitely many pairwise disjoint subsets $\mathcal{R}_1, \ldots, \mathcal{R}_k$ so that $\mathbf{d}^{\mathcal{A}}(w)$ takes the same value for almost all $w \in \mathcal{R}_j$, where $1 \leq j \leq k$. The same result holds also for $\mathbf{c}^{\mathcal{A}}$. (Let us note that k can be larger than 1 even if $\mathcal{M}$ is strongly connected.)

(3) The observations behind the results of (1) and (2) can be used to compute the k and effectively approximate the probability of all $\mathcal{R}_j$ together with the associated values of discrete or timed $\mathcal{A}$-measures up to an arbitrarily small given precision. More precisely, we show that these quantities are expressible using the m-step transition kernel P^m of the product process

$\mathcal{M} \times \mathcal{A}$ defined for $\mathcal{M}$ and $\mathcal{A}$ (see Section 3.2), and we give generic bounds on the number of steps m that is sufficient to achieve the required precision. The m-step transition kernel is defined by nested integrals (see Section 3.1) and can be approximated by numerical methods (see, e.g., [16, 9]). This makes the whole framework effective. The design of more efficient algorithms as well as more detailed analysis applicable to concrete subclasses of SMP are left for future work.

To get some intuition about potential applicability of our results (and about the actual power of DTA which is hidden mainly in their ability to accumulate the total time of several transitions in internal clocks), let us start with a simple example. Consider the following itinerary for travelling between Brno and Prague:

	Brno	Kuřim	Tišnov	Čáslav	Prague
arrival		1:15	2:30	3:30	4:50
departure	0:00	1:20	2:40	3:35	

A traveller has to change a train at each of the three intermediate stops, and she needs at least 3 minutes to walk between the platforms. Assume that all trains depart on time, but can be delayed. Further, assume that travelling time between X and Y has density $f_{X\text{-}Y}$. We wonder what is the chance that a traveller reaches Prague from Brno without missing any train and at most 5 minutes after the scheduled arrival. Answering this question "by hand" is not simple (though still possible). However, it is almost trivial to rephrase this question in terms of DTA measures. The itinerary can be modeled by the following semi-Markov process, where the density f is irrelevant and $\Sigma = \{B,K,T,\check{C},P\}$.
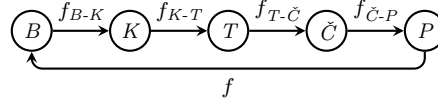
The property of "reaching Prague from Brno without missing any train and at most 5 minutes after the scheduled arrival" is encoded by the DTA $\bar{\mathcal{A}}$ of Figure 1. The automaton uses just one clock x to measure the total elapsed time, and the guards reflect the required timing constraints. Starting in location *init*, the automaton eventually reaches either the location $p\!\uparrow$ or $p\!\downarrow$, which corresponds to satisfaction or violation of the above property, and then it is "restarted". Hence, we are interested in the relative frequency of visits to $p\!\uparrow$ among the visits to $p\!\uparrow$ or $p\!\downarrow$. Using our results, it follows that $\mathbf{d}^{\mathcal{A}}$ is well-defined and takes the same value for almost all runs of $\mathcal{M}$. Hence, the random variable $\mathbf{d}_{p\uparrow}/(\mathbf{d}_{p\uparrow} + \mathbf{d}_{p\downarrow})$ also takes the same value with probability one, and this (unique) value is the quantity of our interest.

Now imagine we wish to model and analyse the flow of passengers in London metro at rush hours. The SMP states then correspond to stations, transition probabilities encode the percentage of passengers traveling in a given direction, and the densities encode the distribution of travelling time. A DTA can be used to monitor a complex list of timing restrictions such as "there is enough time

to change a train", "travelling between important stations does not take more than 30 minutes if one the given routes is used", "trains do not arrive more than 2 minutes later than scheduled", etc. For this we already need several internal clocks. Apart of some auxiliary locations, the constructed DTA would also have special locations used to encode satisfaction/violation of a given restriction (in the DTA $\bar{\mathcal{A}}$ of Figure 1, $(p, \uparrow)$ and $(p, \downarrow)$ are such special locations). Using the results presented in this paper, one may not only study the overall satisfaction of these restrictions, but also estimate the impact of changes in the underlying model (for example, if a given line becomes slower due to some repairs, one may evaluate the decrease in various dependability measures without changing the constructed DTA).

Proof techniques. For a given SMP $\mathcal{M}$ and a given DTA $\mathcal{A}$ we first construct their synchronized product $\mathcal{M} \times \mathcal{A}$, which is another stochastic process. In fact, it turns out that $\mathcal{M} \times \mathcal{A}$ is a discrete-time Markov chain with uncountable state-space. Then, we apply a variant of the standard region construction [2] and thus partition the state-space of $\mathcal{M} \times \mathcal{A}$ into finitely many equivalence classes. At the very core of our paper there are several non-trivial observations about the structure of $\mathcal{M} \times \mathcal{A}$ and its region graph which establish a powerful link to the well-developed ergodic theory of Markov chains with general state-space (see, e.g., [18, 21]). In this way, we obtain the results of items (1) and (2) mentioned above. Some additional work is required to analyze the algorithm presented in Section 4 (whose properties are summarized in item (3) above).

Related work. There is a vast literature on continuous-time Markov chains, semi-Markov processes, or even more general stochastic models such as generalized semi-Markov processes (we refer to, e.g., [23, 6, 20, 17]). In the computer science context, most works on continuous-time stochastic models concern model-checking against a given class of temporal properties [3, 5]. The usefulness of CSL model-checking for dependability analysis is advocated in [14]. Timed automata [2] have been originally used as a model of (non-stochastic) real-time systems. Probabilistic semantics of timed automata is proposed in [4, 7]. The idea of using timed automata as a specification language for continuous-time stochastic processes is relatively recent. In [12], the model-checking problem for continuous-time Markov chains and linear-time properties represented by timed automata is considered (the task is to dermine the probability of all timed words that are accepted by a given timed automaton). A more general model of two-player games over generalized semi-Markov processes with qualitative reachability objectives specified by deterministic timed automata is studied in [10].

2 Preliminaries

In this paper, the sets of all positive integers, non-negative integers, real numbers, positive real numbers, and non-negative real numbers are denoted by $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{R}$, $\mathbb{R}_{>0}$, and $\mathbb{R}_{\geq 0}$, respectively.

Let A be a finite or countably infinite set. A *discrete probability distribution* on A is a function $\alpha : A \to \mathbb{R}_{\geq 0}$ such that $\sum_{a \in A} \alpha(a) = 1$. We say that α is *rational* if $\alpha(a)$ is rational for every $a \in A$. The set of all distributions on A
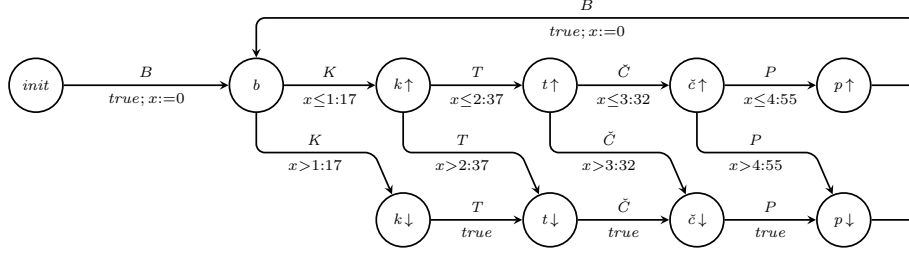
Figure 1: A deterministic timed automaton $\bar{\mathcal{A}}$.

is denoted by $\mathcal{D}(A)$. A *σ-field* over a set Ω is a set $\mathcal{F} \subseteq 2^{\Omega}$ that includes Ω and is closed under complement and countable union. A *measurable space* is a pair $(\Omega, \mathcal{F})$ where Ω is a set called *sample space* and $\mathcal{F}$ is a σ-field over Ω whose elements are called *measurable sets*. A *probability measure* over a measurable space $(\Omega, \mathcal{F})$ is a function $\mathcal{P} : \mathcal{F} \to \mathbb{R}_{\geq 0}$ such that, for each countable collection $\{X_i\}_{i \in I}$ of pairwise disjoint elements of $\mathcal{F}$, $\mathcal{P}(\bigcup_{i \in I} X_i) = \sum_{i \in I} \mathcal{P}(X_i)$, and moreover $\mathcal{P}(\Omega) = 1$. A *probability space* is a triple $(\Omega, \mathcal{F}, \mathcal{P})$, where $(\Omega, \mathcal{F})$ is a measurable space and $\mathcal{P}$ is a probability measure over $(\Omega, \mathcal{F})$. We say that a property $A \subseteq \Omega$ holds *for almost all* elements of a measurable set Y if $\mathcal{P}(Y) > 0$, $A \cap Y \in \mathcal{F}$, and $\mathcal{P}(A \mid Y) = 1$.

All of the integrals used in this paper should be understood as Lebesgue integrals, although we use Riemann-like notation when appropriate.

2.1 Semi-Markov processes

A semi-Markov process (see, e.g., [23]) can be seen as discrete-time Markov chains where each transition is equipped with a density function specifying the distribution of time needed to perform the transition. Formally, let $\mathfrak{D}$ be a set of delay densities, i.e., measurable functions $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$ satisfying $\int_0^{\infty} f(t)\, dt = 1$ where $f(t) = 0$ for every $t < 0$. Moreover, for technical reasons, we assume that each $f \in \mathfrak{D}$ satisfies the following: There is an interval I either of the form $[\ell, u]$ with $\ell, u \in \mathbb{N}_0, \ell < u$, or $[\ell, \infty)$ with $\ell \in \mathbb{N}_0$, such that

- for all $t \in \mathbb{R} \setminus I$ we have that $f(t) = 0$,

- for all $[c, d] \subseteq I$ there is $b > 0$ such that for all $t \in [c, d]$ we have that $f(t) \geq b$.

The assumption that ℓ, u are natural numbers is adopted only for the sake of simplicity. Our results can easily be generalized to the setting where I is an interval with rational bounds or even a finite union of such intervals.

Definition 2.1. *A semi-Markov process (SMP) is a tuple* $\mathcal{M} = (S, \mathbf{P}, \mathbf{D}, \alpha_0)$, *where S is a finite set of states, $\mathbf{P} : S \to \mathcal{D}(S)$ is a transition probability function, $\mathbf{D} : S \times S \to \mathfrak{D}$ is a delay function which to each transition assigns its delay density, and $\alpha_0 \in \mathcal{D}(S)$ is an initial distribution.*

A computation (run) of a SMP $\mathcal{M}$ is initiated in some state s_0, which is chosen randomly according to α_0. In the current state s_i, the next state s_{i+1} is selected randomly according to the distribution $\mathbf{P}(s_i)$, and the selected transition (s_i, s_{i+1}) takes a random time t_i chosen according to the density $\mathbf{D}(s_i, s_{i+1})$. Hence, each run of $\mathcal{M}$ is an infinite timed word $s_0 \, t_0 \, s_1 \, t_1 \cdots$, where $s_i \in S$ and $t_i \in \mathbb{R}_{\geq 0}$ for all $i \in \mathbb{N}_0$. We use $\mathcal{R}_{\mathcal{M}}$ to denote the set of all runs of $\mathcal{M}$.

Now we define a probability space $(\mathcal{R}_{\mathcal{M}}, \mathcal{F}_{\mathcal{M}}, \mathcal{P}_{\mathcal{M}})$ over the runs of $\mathcal{M}$ (we often omit the index $\mathcal{M}$ if it is clear from the context). A *template* is a finite sequence of the form $B = s_0 \, I_0 \, s_1 \, I_1 \cdots s_{n+1}$ such that $n \geq 0$ and I_i is an interval in $\mathbb{R}_{\geq 0}$ for every $0 \leq i \leq n$. Each such B determines the corresponding *cylinder* $\mathcal{R}(B) \subseteq \mathcal{R}$ consisting of all runs of the form $\hat{s}_0 \, t_0 \, \hat{s}_1 \, t_1 \cdots$, where $\hat{s}_i = s_i$ for all $0 \leq i \leq n+1$, and $t_i \in I_i$ for all $0 \leq i \leq n$. The σ-field $\mathcal{F}$ is the Borel σ-field generated by all cylinders. For each template $B = s_0 \, I_0 \, s_1 \, I_1 \cdots s_{n+1}$, let $p_i = \mathbf{P}(s_i)(s_{i+1})$ and $f_i = \mathbf{D}(s_i, s_{i+1})$ for all $0 \leq i \leq n$. The probability $\mathcal{P}(\mathcal{R}(B))$ is defined as follows:

$$\alpha_0(s_0) \cdot \prod_{i=0}^{n} p_i \cdot \int_{t_i \in I_i} f_i(t_i) \, dt_i$$

Then, $\mathcal{P}$ is extended to $\mathcal{F}$ (in the unique way) by applying the extension theorem (see, e.g., [8]).

2.2 Deterministic timed automata

Let $\mathcal{X}$ be a finite set of *clocks*. A *valuation* is a function $\nu : \mathcal{X} \to \mathbb{R}_{\geq 0}$. For every valuation ν and every subset $X \subseteq \mathcal{X}$ of clocks, we use $\nu[X := \mathbf{0}]$ to denote the unique valuation such that $\nu[X := \mathbf{0}](x)$ is equal either to 0 or $\nu(x)$, depending on whether $x \in X$ or not, respectively. Further, for every valuation ν and every $\delta \in \mathbb{R}_{\geq 0}$, the symbol $\nu + \delta$ denotes the unique valuation such that $(\nu + \delta)(x) = \nu(x) + \delta$ for all $x \in \mathcal{X}$. Sometimes we assume an implicit linear ordering on clocks and slightly abuse our notation by identifying a valuation ν with the associated vector of reals.

A *clock constraint* (or *guard*) is a finite conjunction of basic constraints of the form $x \bowtie c$, where $x \in \mathcal{X}$, $\bowtie \in \{<, \leq, >, \geq\}$, and $c \in \mathbb{N}_0$. For every valuation ν and every clock constraint g we have that ν either does or does not satisfy g, written $\nu \models g$ or $\nu \not\models g$, respectively (the satisfaction relation is defined in the expected way). Sometimes we identify a guard g with the set of all valuations that satisfy g and write, e.g., $g \cap g'$. The set of all guards over $\mathcal{X}$ is denoted by $\mathcal{B}(\mathcal{X})$.

Definition 2.2. *A deterministic timed automaton (DTA) is a tuple $\mathcal{A} = (Q, \Sigma, \mathcal{X}, \longrightarrow, q_0)$, where Q is a nonempty finite set of* locations*, Σ is a finite* alphabet*, $\mathcal{X}$ is a finite set of* clocks*, $q_0 \in Q$ is an* initial location*, and $\longrightarrow \subseteq Q \times \Sigma \times \mathcal{B}(\mathcal{X}) \times 2^{\mathcal{X}} \times Q$ is an* edge relation *such that for all $q \in Q$ and $a \in \Sigma$ we have the following:*

1. *the guards are deterministic, i.e., for all edges of the form (q, a, g_1, X_1, q_1) and (q, a, g_2, X_2, q_2) such that $g_1 \cap g_2 \neq \emptyset$ we have that $g_1 = g_2$, $X_1 = X_2$, and $q_1 = q_2$;*

2. *the guards are total, i.e., for all $q \in Q$, $a \in \Sigma$, and every valuation ν there is an edge (q, a, g, X, q') such that $\nu \models g$.*

A *configuration* of $\mathcal{A}$ is a pair (q, ν), where $q \in Q$ and ν is a valuation. An *infinite timed word* over Σ is an infinite sequence $w = c_0 \, c_1 \, c_2 \, c_3 \cdots$, where $c_i \in \Sigma$ when i is even, and $c_i \in \mathbb{R}_{\geq 0}$ when i is odd. The *run* of $\mathcal{A}$ on w is the unique infinite sequence of configurations $\mathcal{A}(w) = (q_0, \nu_0) \, (q_1, \nu_1) \cdots$ such that q_0 is the initial location of $\mathcal{A}$, $\nu_0(x) = 0$ for all $x \in \mathcal{X}$, and for each $i \in \mathbb{N}_0$ we have that

- if c_i is a time stamp, then $q_{i+1} = q_i$ and $\nu_{i+1} = \nu_i + c_i$;

- if c_i is a letter of Σ, then there is a unique edge (q_i, c_i, g, X, q) such that $\nu_i \models g$, and we require that $q_{i+1} = q$ and $\nu_{i+1} = \nu_i[X := \mathbf{0}]$.

Notice that we do not define any acceptance condition for DTA. Instead, we understand DTA as finite-state *observers* that analyze timed words and report about certain events by entering designated locations. The "frequency" of these events is formally captured by the quantities d_q and c_q defined below.

Let $\mathcal{A} = (Q, \Sigma, \mathcal{X}, \longrightarrow, q_0)$ be a DTA, $q \in Q$ some location, and $w = a_0 \, t_0 \, a_1 \, t_1 \cdots$ a timed word over Σ. For every $i \in \mathbb{N}_0$, let $T^i(w)$ be the stamp t_i of w, and $Q^i(w)$ the unique location of $\mathcal{A}$ entered after reading the finite prefix $a_0 \, t_0 \cdots a_i$ of w. Further, let $1_q^i(w)$ be either 1 or 0 depending on whether $Q^i(w) = q$ or not, respectively. The *discrete* and *timed* frequency of visits to q along $\mathcal{A}(w)$, denoted by $\mathbf{d}_q^{\mathcal{A}}(w)$ and $\mathbf{c}_q^{\mathcal{A}}(w)$, are defined in the following way (if $\mathcal{A}$ is clear, it is omitted):

$$
\begin{aligned}
\mathbf{d}_q^{\mathcal{A}}(w) &= \limsup_{n \to \infty} \frac{\sum_{i=1}^{n} 1_q^i(w)}{n} \\
\mathbf{c}_q^{\mathcal{A}}(w) &= \limsup_{n \to \infty} \frac{\sum_{i=1}^{n} T^i(w) \cdot 1_q^i(w)}{\sum_{i=1}^{n} T^i(w)}
\end{aligned}
$$

Hence, every timed word w determines the tuple $\mathbf{d}^{\mathcal{A}} = \left(\mathbf{d}_q^{\mathcal{A}}(w) \right)_{q \in Q}$ and the tuple $\mathbf{c}^{\mathcal{A}} = \left(\mathbf{c}_q^{\mathcal{A}}(w) \right)_{q \in Q}$ of *discrete* and *timed $\mathcal{A}$-measures*, respectively. The $\mathcal{A}$-measures were defined using $\limsup$, because the corresponding limits may not exist in general. If $\lim_{n \to \infty} \sum_{i=1}^{n} 1_q^i(w)/n$ exists for all $q \in Q$, we say that $\mathbf{d}^{\mathcal{A}}$ is *well-defined* for w. Similarly, if $\lim_{n \to \infty} (\sum_{i=1}^{n} T^i(w) \cdot 1_q^i(w))/(\sum_{i=1}^{n} T^i(w))$ exists for all q, we say that $\mathbf{c}^{\mathcal{A}}$ is *well-defined* for w.

As we already noted in Section 1, a DTA $\mathcal{A}$ can be used to observe runs in a given SMP $\mathcal{M}$ after labeling all states of $\mathcal{M}$ with the letters of Σ by a suitable $L : S \to \Sigma$. Then, every run $\sigma = s_0 \, t_0 \, s_1 \, t_1 \cdots$ of $\mathcal{M}$ determines a unique timed word $w_\sigma = L(s_0) \, t_0 \, L(s_1) \, t_1 \cdots$, and one can easily show that for every timed word w, the set $\{\sigma \in \mathcal{R} \mid w_\sigma = w\}$ is measurable in $(\mathcal{R}, \mathcal{F}, \mathcal{P})$.

3 DTA Measures over SMPs

Throughout this section we fix an SMP $\mathcal{M} = (S, \mathbf{P}, \mathbf{D}, \alpha_0)$ and a DTA $\mathcal{A} = (Q, \Sigma, \mathcal{X}, \longrightarrow, q_0)$ where $\mathcal{X} = \{x_1, \ldots, x_n\}$. To simplify our notation, we assume

that $\Sigma = S$, i.e., every run σ of $\mathcal{M}$ is a timed word over Σ (hence, we do not need to introduce any labeling $L : S \rightarrow \Sigma$). This technical assumption does not affect the generality of our results (all of our arguments and proofs work exactly as they are, we only need to rewrite them using less readable notation). Our goal is to prove the following:

Theorem 3.1.

1. $\mathbf{d}^{\mathcal{A}}$ is well-defined for almost all runs of $\mathcal{M}$.

2. There are pairwise disjoint sets $\mathcal{R}_1, \ldots, \mathcal{R}_k$ of runs in $\mathcal{M}$ such that $\mathcal{P}(\mathcal{R}_1 \cup \cdots \cup \mathcal{R}_k) = 1$, and for every $1 \leq j \leq k$ there is a tuple D_j such that $\mathbf{d}^{\mathcal{A}}(\sigma) = D_j$ for almost all $\sigma \in \mathcal{R}_j$ (we use $D_{j,q}$ to denote the q-component of D_j).

In Section 4, we show how to compute the k and approximate $\mathcal{P}(\mathcal{R}_j)$ and D_j up to an arbitrarily small given precision.

An immediate corollary of Theorem 3.1 is an analogous result for $\mathbf{c}^{\mathcal{A}}$.

Corollary 3.2. $\mathbf{c}^{\mathcal{A}}$ is well-defined for almost all runs of $\mathcal{M}$. Further, there are pairwise disjoint sets $\mathcal{R}_1, \ldots, \mathcal{R}_K$ of runs in $\mathcal{M}$ such that $\mathcal{P}(\mathcal{R}_1 \cup \cdots \cup \mathcal{R}_K) = 1$, and for every $1 \leq j \leq K$ there is a tuple C_j such that $\mathbf{c}^{\mathcal{A}}(\sigma) = C_j$ for almost all $\sigma \in \mathcal{R}_j$.

Corollary 3.2 follows from Theorem 3.1 simply by considering the discrete $\mathbf{d}^{S \times \mathcal{A}}$ measure, where the DTA $S \times \mathcal{A}$ is obtained from $\mathcal{A}$ in the following way: the set of locations of $S \times \mathcal{A}$ is $\{q_0\} \cup (S \times Q)$, and for every transition (q_0, s, g, X, q') of $\mathcal{A}$ we add a transition $(q_0, s, g, X, (s, q'))$ to $S \times \mathcal{A}$ and for every transition (q, s, g, X, q') and every $s' \in S$ we add a transition $((s', q), s, g, X, (s, q'))$ to $S \times \mathcal{A}$. The initial location of $S \times \mathcal{A}$ is q_0. Intuitively, $S \times \mathcal{A}$ is the same as $\mathcal{A}$ but it explicitly "remembers" the letter which was used to enter the current location. Let k and D_j be the constants of Theorem 3.1 constructed for $\mathcal{M}$ and $S \times \mathcal{A}$. Observe that the expected time of performing a transition from a given $s \in S$, denoted by E_s, is given by $E_s = \sum_{s' \in S} \mathbf{P}(s)(s') \cdot E_{s,s'}$, where $E_{s,s'}$ is the expectation of a random variable with the density $\mathbf{D}(s, s')$. From this we easily obtain that

$$C_{j,q} = \frac{\sum_{s \in S} E_s \cdot D_{j,(s,q)}}{\sum_{p \in Q} \sum_{s \in S} E_s \cdot D_{j,(s,p)}} \tag{1}$$

for all $q \in Q$ and $1 \leq j \leq k$. The details are given in Appendix A. Hence, we can also compute the constant K and approximate $\mathcal{P}(\mathcal{R}_j)$ and C_j for every $1 \leq j \leq K$ using Equation (1).

It remains to prove Theorem 3.1. Let us start by sketching the overall structure of our proof. First, we construct a synchronous product $\mathcal{M} \times \mathcal{A}$ of $\mathcal{M}$ and $\mathcal{A}$, which is a Markov chain with an uncountable state space $\Gamma_{\mathcal{M} \times \mathcal{A}} = S \times Q \times (\mathbb{R}_{\geq 0})^n$. Intuitively, $\mathcal{M} \times \mathcal{A}$ behaves in the same way as $\mathcal{M}$ and simulates the computation of $\mathcal{A}$ on-the-fly (see Figure 2). Then, we construct a finite region graph $G_{\mathcal{M} \times \mathcal{A}}$ over the product $\mathcal{M} \times \mathcal{A}$. The nodes of $G_{\mathcal{M} \times \mathcal{A}}$ are the sets of states that, roughly speaking, satisfy the same guards of $\mathcal{A}$. Edges are induced by transitions of the product (note that if two states satisfy the same guards, the sets of enabled outgoing transitions are the same). By relying on
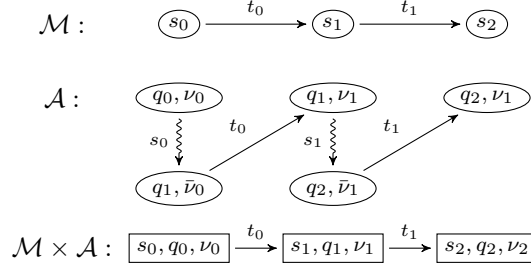
Figure 2: Synchronizing $\mathcal{M}$ and $\mathcal{A}$ in $\mathcal{M} \times \mathcal{A}$. Notice that $\nu_0 = \bar{\nu}_0 = \mathbf{0}$ and $\nu_{i+1} = \bar{\nu}_i + t_i$.

arguments presented in [1, 10], we show that almost all runs reach a node of a bottom strongly connected component (BSCC) $\mathcal{C}$ of $G_{\mathcal{M} \times \mathcal{A}}$ (by definition, each run which enters $\mathcal{C}$ remains in $\mathcal{C}$). This gives us the partition of the set of runs of $\mathcal{M}$ into the sets $\mathcal{R}_1, \ldots, \mathcal{R}_k$ (each $\mathcal{R}_j$ corresponds to one of the BSCCs of $G_{\mathcal{M} \times \mathcal{A}}$).

Subsequently, we concentrate on a fixed BSCC $\mathcal{C}$, and prove that almost all runs that reach $\mathcal{C}$ have the same frequency of visits to a given $q \in Q$ (this gives us the constant $D_{j,q}$). Here we employ several deep results from the theory of general state space Markov chains (see Theorem 3.6). To apply these results, we prove that assuming aperiodicity of $G_{\mathcal{M} \times \mathcal{A}}$ (see Definition 3.10), the state space of the product $\mathcal{M} \times \mathcal{A}$ is *small* (see Definition 3.5 and Lemma 3.11 below). This is perhaps the most demanding part of our proof. Roughly speaking, we show that there is a distinguished subset of states reachable from each state in a fixed number of steps with probability bounded from 0. By applying Theorem 3.6, we obtain a complete invariant distribution on the product, i.e., in principle, we obtain a constant frequency of any non-trivial subset of states. From this we derive our results in a straightforward way. If $G_{\mathcal{M} \times \mathcal{A}}$ is periodic, we use standard techniques for removing periodicity and then basically follow the same stream of arguments as in the aperiodic case.

3.1 General state space Markov chains

We start by recalling the definition of "ordinary" discrete-time Markov chains with discrete state space (DTMC). A DTMC is given by a finite or countably infinite state space S, an initial probability distribution over S, and a one-step transition matrix P which defines the probability $P(s, s')$ of every transion $(s, s') \in S \times S$ so that $\sum_{s' \in S} P(s, s') = 1$ for every $s \in S$. In the setting of uncountable state spaces, transition probabilities cannot be specified by a transition matrix. Instead, one defines the probabilities of moving from a given state s to a given measurable subset X of states. Hence, the concept of transition matrix is replaced with a more general notion of *transition kernel* defined below.

Definition 3.3. *A transition kernel over a measurable space* $(\Gamma, \mathcal{G})$ *is a function* $P : \Gamma \times \mathcal{G} \to [0,1]$ *such that*

    1. $P(z, \cdot)$ is a probability measure over $(\Gamma, \mathcal{G})$ for each $z \in \Gamma$;

    2. $P(\cdot, A)$ is a measurable function for each $A \in \mathcal{G}$ (i.e., for every $c \in \mathbb{R}$, the set of all $z \in \Gamma$ satisfying $P(z, A) \geq c$ belongs to $\mathcal{G}$).

A transition kernel is the core of the following definition.

Definition 3.4. *A* general state space Markov chain (GSSMC) *with a state space* $(\Gamma, \mathcal{G})$*, a transition kernel P and an initial probability measure μ is a stochastic process $\Phi = \Phi_1, \Phi_2, \ldots$ such that each Φ_i is a random variable over a probability space $(\Omega_\Phi, \mathcal{F}_\Phi, \mathcal{P}_\Phi)$ where*

- Ω_Φ *is a set of* runs, *i.e., infinite words over Γ.*

- $\mathcal{F}_\Phi$ *is the product σ-field $\bigotimes_{i=0}^{\infty} \mathcal{G}$.*

- $\mathcal{P}_\Phi$ *is the unique probability measure over $(\Omega_\Phi, \mathcal{F}_\Phi)$ such that for every finite sequence $A_0, \cdots, A_n \in \mathcal{F}_\Phi$ we have that $\mathcal{P}_\Phi(\Phi_0 \in A_0, \cdots, \Phi_n \in A_n)$ is equal to*

$$\int_{y_0 \in A_0} \cdots \int_{y_{n-1} \in A_{n-1}} \mu(dy_0) \cdot P(y_0, dy_1) \cdots P(y_{n-1}, A_n). \qquad (2)$$

- *Each Φ_i is the projection of elements of Ω_Φ onto the i-th component.*

A *path* is a finite sequence $z_1 \cdots z_n$ of states from Γ. From Equation (2) we get that Φ also satisfies the following properties which will be used to show several results about the chain Φ by working with the transition kernel only.

    1. $\mathcal{P}_\Phi(\Phi_0 \in A_0) = \mu(A_0)$,

    2. $\mathcal{P}_\Phi(\Phi_{n+1} \in A \mid \Phi_n, \ldots, \Phi_0) = \mathcal{P}_\Phi(\Phi_{n+1} \in A \mid \Phi_n) = P(\Phi_n, A)$ almost surely,

    3. $\mathcal{P}_\Phi(\Phi_{n+m} \in A \mid \Phi_n) = P^m(\Phi_n, A)$ almost surely,

where the m-step transition kernel P^m is defined as follows:

$$P^1(z, A) = P(z, A)$$

$$P^{i+1}(z, A) = \int_\Gamma P(z, dy) \cdot P^i(y, A).$$

Notice that the transition kernel and the m-step transition kernel are analogous counterparts to the transition matrix and the k-step transition matrix of a DTMC.

As we mentioned above, our proof of Theorem 3.1 employs several results of GSSMC theory. In particular, we make use of the notion of *smallness* of the state space defined as follows.

Definition 3.5. *Let $m \in \mathbb{N}$, $\varepsilon > 0$, and ν be a probability measure on $\mathcal{G}$. A set $C \in \mathcal{G}$ is (m, ε, ν)-small if for all $x \in C$ and $B \in \mathcal{G}$ we have that $P^m(x, B) \geq \varepsilon \cdot \nu(B)$.*

GSSMCs where the whole state space is small have many nice properties, and the relevant ones are summarized in the following theorem.

Theorem 3.6. *If Γ is (m, ε, ν)-small, then*

1. [**Existence of invariant measure**] *There exists a unique probability measure π such that for all $A \in \mathcal{G}$ we have that*

$$\pi(A) \quad = \quad \int_\Gamma \pi(dx) P(x, A)$$

2. [**Strong law of large numbers**] *If $h : \Gamma \to \mathbb{R}$ satisfies $\int_\Gamma h(x)\pi(dx) < \infty$, then almost surely*

$$\lim_{n \to \infty} \frac{\sum_{i=1}^{n} h(\Phi_i)}{n} \quad = \quad \int_\Gamma h(x)\pi(dx)$$

3. [**Uniform ergodicity**] *For all $x \in \Gamma$, $A \in \mathcal{G}$, and all $n \in \mathbb{N}$,*

$$\sup_{A \in \mathcal{G}} |P^n(x, A) - \pi(A)| \quad \leq \quad (1 - \varepsilon)^{\lfloor n/m \rfloor}$$

Proof. The theorem is a consequence of stadard results for GSSMCs. Since Γ is (m, ε, ν)-small, we have

(i) Φ is by definition φ-irreducible for $\varphi = \nu$, and thus also ψ-irreducible by [18, Proposition 4.2.2];

(ii) Γ is by definition also (a, ε, ν)-petite (see [18, Section 5.5.2]), where a is the Dirac distribution on $\mathbb{N}_0$ with $a(m) = 1$, $a(n) = 0$ for $n \neq m$;

(iii) the first return time to Γ is trivially 1.

ad 1. By (iii), Γ is not uniformly transient, hence by (i), (ii) and [18, Theorem 8.0.2], Φ is recurrent. Thus by [18, Theorem 10.0.1], there exists a unique invariant probability measure π.

ad 2. By (i)-(iii) and [18, Theorem 10.4.10 (ii)], Φ is positive Harris. Therefore, we may apply [18, Theorem 17.0.1 (i)] and obtain the desired result.

ad 3. This follows immediately from [21, Theorem 8]. $\qquad\qquad\square$

3.2 The product process

The *product process* of $\mathcal{M}$ and $\mathcal{A}$, denoted by $\mathcal{M} \times \mathcal{A}$, is a GSSMC with the state space $\Gamma_{\mathcal{M} \times \mathcal{A}} = S \times Q \times (\mathbb{R}_{\geq 0})^n$, where $n = |\mathcal{X}|$ is the number of clocks of $\mathcal{A}$. The σ-field over $\Gamma_{\mathcal{M} \times \mathcal{A}}$ is the product σ-field $\mathcal{G}_{\mathcal{M} \times \mathcal{A}} = 2^S \otimes 2^Q \otimes \mathfrak{B}^n$ where $\mathfrak{B}^n$ is the Borel σ-field over the set $(\mathbb{R}_{\geq 0})^n$. For each $A \in \mathcal{G}_{\mathcal{M} \times \mathcal{A}}$, the initial probability $\mu_{\mathcal{M} \times \mathcal{A}}(A)$ is equal to $\sum_{(s,q_0,\mathbf{0}) \in A} \alpha_0(s)$ (recall that α_0 is the initial distribution of $\mathcal{M}$).

The behavior of $\mathcal{M} \times \mathcal{A}$ is depicted in Figure 2. Each step of the product process corresponds to one step of $\mathcal{M}$ and two steps of $\mathcal{A}$. The step of the product starts by simulating the discrete step of $\mathcal{A}$ that reads the current state of $\mathcal{M}$ and possibly resets some clocks, followed by simulating simultaneously the step of $\mathcal{M}$ that takes time t and the corresponding step of $\mathcal{A}$ which reads the time stamp t.

Now we define the transition kernel $P_{\mathcal{M} \times \mathcal{A}}$ of the product process. Let $z = (s, q, \nu)$ be a state of $\Gamma_{\mathcal{M} \times \mathcal{A}}$, and let $(\bar{q}, \bar{\nu})$ be the configuration of $\mathcal{A}$ entered from the configuration (q, ν) after reading s (note that $\bar{\nu}$ is not necessarily the same as ν because $\mathcal{A}$ may reset some clocks). It suffices to define $P_{\mathcal{M} \times \mathcal{A}}(z, \cdot)$ only for generators of $\mathcal{G}_{\mathcal{M} \times \mathcal{A}}$ and then apply the extension theorem (see, e.g., [8]) to obtain a unique probability measure $P_{\mathcal{M} \times \mathcal{A}}(z, \cdot)$ over $(\Gamma_{\mathcal{M} \times \mathcal{A}}, \mathcal{G}_{\mathcal{M} \times \mathcal{A}})$. Generators of $\mathcal{G}_{\mathcal{M} \times \mathcal{A}}$ are sets of the form $\{s'\} \times \{q'\} \times \mathbf{I}$ where $s' \in S$, $q' \in Q$ and $\mathbf{I}$ is the product $I_1 \times \cdots \times I_n$ of intervals I_i in $\mathbb{R}_{\geq 0}$. If $q' \neq \bar{q}$, then we define $P_{\mathcal{M} \times \mathcal{A}}(z, \{s'\} \times \{q'\} \times \mathbf{I}) = 0$. Otherwise, we define

$$P_{\mathcal{M} \times \mathcal{A}}(z, \{s'\} \times \{q'\} \times \mathbf{I}) = \mathbf{P}(s)(s') \cdot \int_0^\infty f(t) \cdot 1_{\mathbf{I}}(\bar{\nu} + t) dt$$

Here $f = \mathbf{D}(s, s')$ and $1_{\mathbf{I}}$ is the indicator function of the set $\mathbf{I}$.

Since $P_{\mathcal{M} \times \mathcal{A}}(z, \cdot)$ is by definition a probability measure over $(\Gamma_{\mathcal{M} \times \mathcal{A}}, \mathcal{G}_{\mathcal{M} \times \mathcal{A}})$, it remains to check the second condition of Definition 3.3.

Lemma 3.7. *Let $A \in \mathcal{G}_{\mathcal{M} \times \mathcal{A}}$. Then $P_{\mathcal{M} \times \mathcal{A}}(\cdot, A)$ is a measurable function, i.e., $\mathcal{M} \times \mathcal{A}$ is a GSSMC.*

A proof of this lemma can be found in Appendix B.1. Recall that by Definition 3.4, $\mathcal{P}_{\mathcal{M} \times \mathcal{A}}$ is the unique probability measure on the product σ-field $\mathcal{F}_{\mathcal{M} \times \mathcal{A}} = \bigotimes_{i=0}^\infty \mathcal{G}_{\mathcal{M} \times \mathcal{A}}$ induced by $P_{\mathcal{M} \times \mathcal{A}}$ and the initial probability measure $\mu_{\mathcal{M} \times \mathcal{A}}$.

3.2.1 The correspondence between $\mathcal{M} \times \mathcal{A}$ and $\mathcal{M}$

In this subsection we show that $\mathcal{M} \times \mathcal{A}$ correctly reflects the behaviour of $\mathcal{M}$. First, we define the $\mathbf{d}^{\mathcal{A}}$ measure for $\mathcal{M} \times \mathcal{A}$. (As the DTA $\mathcal{A}$ is fixed, we omit them and write $\mathbf{d}$ and $\mathbf{d}_q$ instead of $\mathbf{d}^{\mathcal{A}}$ and $\mathbf{d}_q^{\mathcal{A}}$, respectively.) Let $\sigma = (s_0, q_0, \nu_0)(s_1, q_1, \nu_1) \cdots$ be a run of $\mathcal{M} \times \mathcal{A}$ and $q \in Q$ a location. For every $i \in \mathbb{N}_0$, let $1_q^i(\sigma)$ be either 1 or 0 depending on whether if $q_i = q$ or not, respectively. We put

$$\mathbf{d}_q(\sigma) = \limsup_{n \to \infty} \frac{\sum_{i=1}^n 1_q^i(\sigma)}{n}$$

14

Lemma 3.8. *There is a measurable one-to-one mapping ξ from the set of runs of $\mathcal{M}$ to the set of runs of $\mathcal{M} \times \mathcal{A}$ such that*

- *ξ preserves measure, i.e., for every measurable set X of runs of $\mathcal{M}$ we have that $\xi(X)$ is also measurable and $\mathcal{P}_{\mathcal{M}}(X) = \mathcal{P}_{\mathcal{M} \times \mathcal{A}}(\xi(X))$;*

- *ξ preserves $\mathbf{d}$, i.e., for every run σ of $\mathcal{M}$ and every $q \in Q$ we have that $\mathbf{d}_q(\sigma)$ is well-defined iff $\mathbf{d}_q(\xi(\sigma))$ is well-defined, and $\mathbf{d}_q(\sigma) = \mathbf{d}_q(\xi(\sigma))$.*

A formal proof of Lemma 3.8 is given in Appendix B.2.

3.2.2 The region graph of $\mathcal{M} \times \mathcal{A}$

Although the state-space $\Gamma_{\mathcal{M} \times \mathcal{A}}$ is uncountable, we can define the standard *region relation* $\sim$ [2] over $\Gamma_{\mathcal{M} \times \mathcal{A}}$ with finite index, and then work with finitely many *regions*. For a given $a \in \mathbb{R}$, we use $frac(a)$ to denote the fractional part of a, and $int(a)$ to denote the integral part of a. For $a, b \in \mathbb{R}$, we say that a and b *agree on integral part* if $int(a) = int(b)$ and neither or both a, b are integers.

We denote by $B_{\max}$ the maximal constant that appears in the guards of $\mathcal{A}$ and say that a clock $x \in \mathcal{X}$ is *relevant* for ν if $\nu(x) \leq B_{\max}$. Finally, we put $(s_1, q_1, \nu_1) \sim (s_2, q_2, \nu_2)$ if

- $s_1 = s_2$ and $q_1 = q_2$;

- for all relevant $x \in \mathcal{X}$ we have that $\nu_1(x)$ and $\nu_2(x)$ agree on integral parts;

- for all relevant $x, y \in \mathcal{X}$ we have that $frac(\nu_1(x)) \leq frac(\nu_1(y))$ iff $frac(\nu_2(x)) \leq frac(\nu_2(y))$.

Note that $\sim$ is an equivalence with finite index. The equivalence classes of $\sim$ are called *regions*. Observe that states in the same region have the same behavior with respect to qualitative reachability. This is formalized in the following lemma.

Lemma 3.9. *Let R and T be regions and $z, z' \in R$. Then $P_{\mathcal{M} \times \mathcal{A}}(z, T) > 0$ iff $P_{\mathcal{M} \times \mathcal{A}}(z', T) > 0$.*

A proof of Lemma 3.9 can be found in [10]. Further, we define a finite *region graph* $G_{\mathcal{M} \times \mathcal{A}} = (V, E)$ where the set of vertices V is the set of regions and for every pair of regions R, R' there is an edge $(R, R') \in E$ iff $P_{\mathcal{M} \times \mathcal{A}}(z, R') > 0$ for some $z \in R$ (due to Lemma 3.9, the concrete choice of z is irrelevant). For technical reasons, we assume that V contains only regions reachable with positive probability in $\mathcal{M} \times \mathcal{A}$.

3.3 Finishing the proof of Theorem 3.1

Our proof is divided into three parts. In the first part we consider a general region graph which is not necessarily strongly connected, and show that we can actually concentrate just on its BSCCs. In the second part we study a given BSCC under the aperiodicity assumption. Finally, in the last part we consider a general BSCC which may be periodic. (The second part is included mainly for the sake of readability.)

Non-strongly connected region graph

Let $C_1, \ldots, C_k$ be the BSCCs of the region graph. The set $\mathcal{R}_i$ consists of all runs σ of $\mathcal{M}$ such that $\xi(\omega)$ visits (a configuration in a region of) C_i, where ξ is the mapping of Lemma 3.8. By applying the arguments of [1, 10], it follows that almost runs in $\mathcal{M} \times \mathcal{A}$ visit a configuration of a BSCC. By Lemma 3.8, ξ preserves $\mathbf{d}$ and the probability $\mathcal{P}_{\mathcal{M}}(\mathcal{R}_i)$ is equal to the probability of visiting C_i in $\mathcal{M} \times \mathcal{A}$. Further, since the value of $\mathbf{d}$ does not depend on a finite prefix of a run, we may safely assume that $\mathcal{M} \times \mathcal{A}$ is initialized in C_i in such a way that the initial distribution corresponds to the conditional distribution of the first visit to C_i conditioned on visiting C_i.

In a BSCC C_i, there may be some *growing* clocks that are never reset. Since the values of growing clocks are just constantly increasing, the product process never returns to a state it has visited before. Therefore, there is no invariant distribution. Observe that all runs initiated in C_i eventually reach a configuration where the values of all growing clocks are larger than the maximal constant $B_{\max}$ employed in the guards of $\mathcal{A}$. This means that C_i actually consists *only* of regions where all growing clocks are irrelevant (see Section 3.2.2), because C_i would not be strongly connected otherwise. Hence, we can safely *remove* every growing clock x from C_i, replacing all guards of the form $x > c$ or $x \geq c$ with *true* and all guards of the form $x < c$ or $x \leq c$ with *false*. So, from now on we assume that there are no growing clocks in C_i.

Strongly connected & aperiodic region graph

In this part we consider a given BSCC C_i of the region graph $G_{\mathcal{M} \times \mathcal{A}}$. This is equivalent to assuming that $G_{\mathcal{M} \times \mathcal{A}}$ is strongly connected and $\Gamma_{\mathcal{M} \times \mathcal{A}}$ is equal to the union of all regions of $G_{\mathcal{M} \times \mathcal{A}}$ (recall that $G_{\mathcal{M} \times \mathcal{A}}$ consists just of regions reachable with positive probability in $\mathcal{M} \times \mathcal{A}$). We also assume that there are no growing clocks (see the previous part). Further, in this subsection we assume that $G_{\mathcal{M} \times \mathcal{A}}$ is aperiodic in the following sense.

Definition 3.10. *A period p of the region graph $G_{\mathcal{M} \times \mathcal{A}}$ is the greatest common divisor of lengths of all cycles in $G_{\mathcal{M} \times \mathcal{A}}$. The region graph $G_{\mathcal{M} \times \mathcal{A}}$ is aperiodic if $p = 1$.*

The key to proving Theorem 3.1 in the current restricted setting is to show that the state space of $\mathcal{M} \times \mathcal{A}$ is small (recall Definition 3.5) and then apply Theorem 3.6 (1) and (2) to obtain the required characterization of the long-run behavior of $\mathcal{M} \times \mathcal{A}$.

Proposition 3.11. *Assume that $G_{\mathcal{M} \times \mathcal{A}}$ is strongly connected and aperiodic. Then there exist a region R, a measurable subset $S \subseteq R$, $n \in \mathbb{N}$, $b > 0$, and a probability measure κ such that $\kappa(S) = 1$ and for all measurable $T \subseteq S$ and $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}$ we have that $P^n_{\mathcal{M} \times \mathcal{A}}(z, T) > b \cdot \kappa(T)$. In other words, the set $\Gamma_{\mathcal{M} \times \mathcal{A}}$ of all states of the GSSMC $\mathcal{M} \times \mathcal{A}$ is (n, b, κ)-small.*

Sketch. We show that there exist $z^* \in \Gamma_{\mathcal{M} \times \mathcal{A}}$, $n \in \mathbb{N}$, and $\gamma > 0$ such that for an arbitrary starting state $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}$ there is a path from z to z^* of length

exactly n that is γ-*wide* in the sense that the waiting time of any transition in the path can be changed by $\pm\gamma$ without ending up in a different region in the end. The target set S then corresponds to a "neighbourhood" of z^* within the region of z^*. Any small enough sub-neighbourhood of z^* is visited by a set of runs that follow the γ-wide path closely enough. The probability of this set of runs then depends linearly on the size of the sub-neighbourhood when measured by κ, where κ is essentially the Lebesgue measure restricted to S.

So, it remains to find suitable z^*, n, and γ. For a given starting state $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}$, we construct a path of fixed length n (independent of z) that always ends in the same state z^*. Further, the path is γ-wide for some $\gamma > 0$ independent of z. Technically, the path is obtained by concatenating five sub-paths each of which has a fixed length independent of z. These sub-paths are described in greater detail below.

In the first sub-path, we move to a δ-*separated* state for some fixed $\delta > 0$ independent of z. A state is δ-separated if the fractional parts of all relevant clocks are approximately equally distributed on the $[0, 1]$ line segment (each two of them have distance at least δ). We can easily build the first sub-path so that it is δ-wide.

For the second sub-path, we first fix some region R_1. Since $G_{\mathcal{M} \times \mathcal{A}}$ is strongly connected and aperiodic, there is a fixed n' such that R_1 is reachable from an arbitrary state of $\Gamma_{\mathcal{M} \times \mathcal{A}}$ in exactly n' transitions. The second sub-path is chosen as a (δ/n')-wide path of length n' that leads to a (δ/n')-separated state of R_1 (we show that such a sub-path is guaranteed to exist; intuitively, the reason why the separation and wideness may decrease proportionally to n' is that the fractional parts of relevant clock may be forced to move closer and closer to each other by the resets performed along the sub-path).

In the third sub-path, we squeeze the fractional parts of all relevant clocks close to 0. We go through a fixed region path $R_1 \cdots R_k$ (independent of z) so that in each step we shift the time by an integral value minus a small constant c (note that the fractional parts of clocks reset during this path have fixed relative distances). Thus, we reach a state z'_k that is "almost fixed" in the sense that the values of all relevant clocks in z'_k are the same for every starting state z. Note that the third sub-path is c-wide. At this point, we should note that if we defined the product process somewhat differently by identifying all states differing only in the values of irrelevant clocks (which does not lead to any technical complications), we would be done, i.e., we could put $z^* = z'_k$. We have neglected this possibility mainly for presentation reasons. So, we need two more sub-paths to fix the values of irrelevant clocks.

In the fourth sub-path, we act similarly as in the first sub-path and prepare ourselves for the final sub-path. We reach a δ-separated state that is almost equal to a fixed state $z_\ell \in R_\ell$. Again, we do it by a δ-wide path of a fixed length.

In the fifth sub-path, we follow a fixed region path $R_\ell \cdots R_{\ell+m}$ such that each clock not relevant in R_ℓ is reset along this path, and hence we reach a fixed state $z^* \in R_{\ell+m}$. Here we use our assumption that every clock can be reset to zero (i.e., there are no growing clocks). $\qquad\square$

Now we may finish the proof of Theorem 3.1. By Theorem 3.6 (1), there is a

unique invariant distribution π on $\Gamma_{\mathcal{M} \times \mathcal{A}}$. For every $q \in Q$, we denote by A_q the set of all states of $\mathcal{M} \times \mathcal{A}$ of the form $(s, q, \nu) \in \Gamma_{\mathcal{M} \times \mathcal{A}}$. By Theorem 3.6 (2), for almost all runs σ of $\mathcal{M} \times \mathcal{A}$ we have that $\mathbf{d}(\sigma)$ is well-defined and $\mathbf{d}_q(\sigma) = \sum \pi(A_q)$. By Lemma 3.8, we obtain the same for almost all runs of $\mathcal{M}$.

Strongly connected & periodic region graph

Now we consider a general BSCC $\mathcal{C}_i$ of the region graph $G_{\mathcal{M} \times \mathcal{A}}$. Technically, we adopt the same setup as the previous part but remove the aperiodicity condition. That is, we assume that $G_{\mathcal{M} \times \mathcal{A}}$ is strongly connected, $\Gamma_{\mathcal{M} \times \mathcal{A}}$ is equal to the union of all regions of $G_{\mathcal{M} \times \mathcal{A}}$, and there are no growing clocks.

Let p be the period of $G_{\mathcal{M} \times \mathcal{A}}$. In this case, $\mathcal{M} \times \mathcal{A}$ is not necessarily small in the sense of Definition 3.5. By employing standard methods for periodic Markov chains, we decompose $\mathcal{M} \times \mathcal{A}$ into p stochastic processes $\Phi_0, \ldots, \Phi_{p-1}$ where each Φ_k makes steps corresponding to p steps of the original process $\mathcal{M} \times \mathcal{A}$ (except for the first step which corresponds just to k steps of $\mathcal{M} \times \mathcal{A}$). Each Φ_k is aperiodic and hence small (this follows by slightly generalizing the arguments of the previous part; see Proposition 3.13). Thus, we can apply Theorem 3.6 to each Φ_k separately and express the frequency of visits to q in Φ_k in terms of a unique invariant distribution π_k for Φ_k. Finally, we obtain the frequency of visits to q in $\mathcal{M} \times \mathcal{A}$ as an average of the corresponding frequencies in Φ_k.

Let us start by decomposing the set of nodes V of $G_{\mathcal{M} \times \mathcal{A}}$ into p classes that constitute a cyclic structure (see e.g. [11, Theorem 4.1]).

Lemma 3.12. *There are disjoint sets $V_0, \ldots, V_{p-1} \subseteq V$ such that $V = \bigcup_{k=0}^{p-1} V_k$ and for all $u, v \in V$ we have that $(u, v) \in E$ iff there is $k \in \{0, \ldots, p-1\}$ satisfying $u \in V_k$ and $v \in V_j$ where $j = (k+1) \mod p$.*

For each $k \in \{0, \ldots, p-1\}$ we construct a GSSMC Φ_k with state space $\Gamma_{\mathcal{M} \times \mathcal{A}}^k = \bigcup_{R \in V_k} R$, a transition kernel $P^p(\cdot, \cdot)$ restricted to $\Gamma_{\mathcal{M} \times \mathcal{A}}^k$, and an initial probability measure μ_k defined by $\mu_k(A) = \int_{z \in \Gamma_{\mathcal{M} \times \mathcal{A}}} \mu(dz) \cdot P^k(z, A)$. For each k, we define the discrete frequency $\mathbf{d}_q^k$ of visits q in the process Φ_k. Then we show that if $\mathbf{d}^k$ is well-defined in Φ_k, we can express the frequency $\mathbf{d}_q$ in $\mathcal{M} \times \mathcal{A}$.

Note that for every run $z_0 z_1 \cdots$ of $\mathcal{M} \times \mathcal{A}$, the word $z_k z_{p+k} z_{2p+k}$ is a run of Φ_k. For a run $\sigma = (s_0, q_0, \nu_0)(s_1, q_1, \nu_1) \cdots$, $k \in \{0, \ldots, p-1\}$, and a location $q \in Q$, let define $1_q^{i,k}(\sigma)$ to be either 1 or 0 depending on whether $q_{ip+k} = q$ or not, respectively. Further, we put

$$\mathbf{d}_q^k(\sigma) = \limsup_{n \to \infty} \frac{\sum_{i=1}^n 1_q^{i,k}(\sigma)}{n}$$

Assuming that each $\mathbf{d}^k$ is well-defined, for almost all runs σ of $\mathcal{M} \times \mathcal{A}$ we have the following:

$$
\begin{aligned}
\mathbf{d}_q(\sigma) &= \lim_{n \to \infty} \frac{\sum_{i=1}^n 1_q^i(\sigma)}{n} = \lim_{n \to \infty} \frac{\sum_{i=1}^n \sum_{k=0}^{p-1} 1_q^{i,k}(\sigma)}{np} \\
&= \frac{1}{p} \sum_{k=0}^{p-1} \lim_{n \to \infty} \frac{\sum_{i=1}^n 1_q^{i,k}(\sigma)}{n} = \frac{1}{p} \sum_{k=0}^{p-1} \mathbf{d}_q^k(\sigma)
\end{aligned}
$$

18

So, it suffices to concentrate on $\mathbf{d}_q^k$. The following proposition is a generalization of Proposition 3.11 to periodic processes.

Proposition 3.13. *Assume that $G_{\mathcal{M} \times \mathcal{A}}$ is strongly connected and has a period p. For every $k \in \{0, \ldots, p-1\}$ there exist a region $R_k \in V_k$, a measurable $S_k \subset R_k$, $n_k \in \mathbb{N}$, $b_k > 0$, and a probability measure κ_k such that $\kappa_k(S_k) = 1$ and for every measurable $T \subseteq S_k$ and $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}^k$ we have $P_{\mathcal{M} \times \mathcal{A}}^{n_k \cdot p}(z, T) > b_k \cdot \kappa_k(T)$. In other words, Φ_k is (n_k, b_k, κ_k)-small.*

By Theorem 3.6 (1), for every $k \in \{0, \ldots, p-1\}$, there is a unique invariant distribution π_k on $\Gamma_{\mathcal{M} \times \mathcal{A}}$ for the process Φ_k. By Theorem 3.6 (2), each $\mathbf{d}^k$ is well-defined and for almost all runs σ we have that $\mathbf{d}_q^k(\sigma) = \pi_k(A_q)$. Thus, we obtain

$$\mathbf{d}_q(\sigma) = \frac{1}{p} \sum_{k=0}^{p-1} \pi_k(A_q)$$

4   Approximating DTA Measures

In this section we show how to approximate the DTA measures for SMPs using the m-step transition kernel $P_{\mathcal{M} \times \mathcal{A}}^m$ of $\mathcal{M} \times \mathcal{A}$. The procedure for computing $P_{\mathcal{M} \times \mathcal{A}}^m$ up to a sufficient precision is taken as a "black box" part of the algorithm, we concentrate just on developing generic bounds on m that are sufficient to achieve the required precision.

For simplicity, we assume that the initial distribution α_0 of $\mathcal{M}$ assigns 1 to some $s_0 \in S$ (all of the results presented in this section can easily be generalized to an arbitrary initial distribution). The initial state in $\mathcal{M} \times \mathcal{A}$ is $z_0 = (s_0, q_0, \mathbf{0})$.

As we already noted in the previous section, the constant k of Theorem 3.1 is the number of BSCCs of $G_{\mathcal{M} \times \mathcal{A}}$. For the rest of this section, we fix some $1 \leq j \leq k$, and write just $\mathcal{C}$, $\mathcal{R}$ and D instead of $\mathcal{C}_j$, $\mathcal{R}_j$ and D_j, respectively. We slightly abuse our notation by using $\mathcal{C}$ to denote also the set of configurations that belong to some region of $\mathcal{C}$ (particularly in expressions such as $P_{\mathcal{M} \times \mathcal{A}}(z, \mathcal{C})$).

The probability $\mathcal{P}_{\mathcal{M}}(\mathcal{R})$ is equal to the probability of visiting $\mathcal{C}$ in $\mathcal{M} \times \mathcal{A}$. Observe that

$$\mathcal{P}_{\mathcal{M}}(\mathcal{R}) = \lim_{i \to \infty} P_{\mathcal{M} \times \mathcal{A}}^i(z_0, \mathcal{C})$$

Let us analyze the speed of this approximation. First, we need to introduce several parameters. Let $p_{\min}$ be the smallest transition probability in $\mathcal{M}$, and $\mathfrak{D}(\mathcal{M})$ the set of delay densities used in $\mathcal{M}$, i.e., $\mathfrak{D}(\mathcal{M}) = \{\mathbf{D}(s, s') \mid s, s' \in S\}$. Let $|V|$ be the number of vertices (regions) of $G_{\mathcal{M} \times \mathcal{A}}$. Due to our assumptions imposed on delay densities, there is a fixed bound $c_{\mathfrak{D}} > 0$ such that, for all $f \in \mathfrak{D}(\mathcal{M})$ and $x \in [0, B_{\max}]$, either $f(x) > c_{\mathfrak{D}}$ or $f(x) = 0$. Further, $\int_{B_{\max}}^{\infty} f(x)dx$ is either larger than $c_{\mathfrak{D}}$ or equal to 0.

Theorem 4.1. *For every $i \in \mathbb{N}$ we have that*

$$\mathcal{P}_{\mathcal{M}}(\mathcal{R}) - P_{\mathcal{M} \times \mathcal{A}}^i(z_0, \mathcal{C}) \quad \leq \quad \left(1 - \left(\frac{p_{min} \cdot c_{\mathfrak{D}}}{c}\right)^c\right)^{\lfloor i/c \rfloor}$$

where $c = 4 \cdot |V|$.

Sketch. We denote by B the union of all regions that belong to BSCCs of $G_{\mathcal{M}\times\mathcal{A}}$. We show that for $c = 4 \cdot |V|$ there is a lower bound $p_{bound} = (p_{\min} \cdot c_{\mathfrak{D}} \cdot 1/c)^c$ on the probability of reaching B in at most c steps from any state $z \in \Gamma_{\mathcal{M}\times\mathcal{A}}$. Note that then the probability of not hitting B after $i = m \cdot c$ steps is at most $(1 - p_{bound})^m$. However, this means that $P^i_{\mathcal{M}\times\mathcal{A}}(z, \mathcal{C})$ cannot differ from the probability of reaching $\mathcal{C}$ (and thus also from $\mathcal{P}_{\mathcal{M}}(\mathcal{R})$) by more than $(1 - p_{bound})^m$ because $\mathcal{C} \subseteq B$ and the probability of reaching $\mathcal{C}$ from $B \smallsetminus \mathcal{C}$ is 0.

The bound p_{bound} is provided by arguments similar to the proof of Proposition 3.11. From any state z we build a δ-wide path to a state in B that has length bounded by $4 \cdot |V|$ such that $\delta = p_{\min} \cdot c_{\mathfrak{D}} \cdot 1/c$. The paths that follow this δ-wide path closely enough (hence, reach B) have probability p_{bound}. □

Now let us concentrate on approximating the tuple D. This can be done by considering just the BSCC $\mathcal{C}$. Similarly as in Section 3, from now on we assume that $\mathcal{C}$ is the set of nodes of $G_{\mathcal{M}\times\mathcal{A}}$ (i.e., $G_{\mathcal{M}\times\mathcal{A}}$ is strongly-connected) and that $\Gamma_{\mathcal{M}\times\mathcal{A}}$ is equal to the union of all regions of $\mathcal{C}$.

As in Section 3, we start with the aperiodic case. Then, Theorem 3.6 (3.) implies that each D_q can be approximated using $P^i_{\mathcal{M}\times\mathcal{A}}(u, A_q)$ where u is an arbitrary state of $\Gamma_{\mathcal{M}\times\mathcal{A}}$ and A_q is the set of all states of $\mathcal{M} \times \mathcal{A}$ of the form (s, q, ν). More precisely, we obtain the following:

Theorem 4.2. *Assume that $G_{\mathcal{M}\times\mathcal{A}}$ is strongly connected and aperiodic. Then for all $i \in \mathbb{N}$, $u \in \Gamma_{\mathcal{M}\times\mathcal{A}}$, and $q \in Q$*

$$\left| D_q - P^i_{\mathcal{M}\times\mathcal{A}}(u, A_q) \right| \leq \left(1 - \left(\frac{p_{\min} \cdot c_{\mathfrak{D}}}{r} \right)^r \right)^{\lfloor i/r \rfloor}$$

where $r = \lfloor |V|^{4 \ln |V|} \rfloor$.

Proof. From the proof of Proposition 3.13 (for details see Appendix C), we obtain that $\Gamma_{\mathcal{M}\times\mathcal{A}}$ is (m, ε, κ)-small with $m \leq r$ and $\varepsilon = (\frac{p_{\min}c_{\mathfrak{D}}}{r})^r$, and the result follows from Theorem 3.6 (3.). □

Now let us consider the general (periodic) case. We adopt the same notation as in Section 3, i.e., the period of $G_{\mathcal{M}\times\mathcal{A}}$ is denoted by p, the decomposition of the set V by $V_0, \ldots, V_{p-1}$ (see Lemma 3.12), and $\Gamma^k_{\mathcal{M}\times\mathcal{A}}$ denotes the set $\bigcup_{R \in V_k} R$ for every $k \in \{0, \ldots, p-1\}$.

Theorem 4.3. *For every $i \in \mathbb{N}$ we have that*

$$\left| D_q - \frac{1}{p} \cdot \sum_{k=0}^{p-1} P^{i \cdot p}_{\mathcal{M}\times\mathcal{A}}(u_k, A_q) \right| \leq \left(1 - \left(\frac{p_{\min} \cdot c_{\mathfrak{D}}}{r} \right)^r \right)^{\lfloor i/r \rfloor}$$

where $u_k \in \Gamma^k_{\mathcal{M}\times\mathcal{A}}$ and $r = \lfloor |V|^{4 \ln |V|} \rfloor$.

Proof. Due to the results of Section 3 we have that $D_q = \frac{1}{p} \cdot \sum_{k=0}^{p-1} \pi_k(A_q)$, where π_k is the invariant measure for the k-th aperiodic decomposition Φ_k of the product process $\mathcal{M} \times \mathcal{A}$ (i.e. π_k is a measure over $\Gamma^k_{\mathcal{M}\times\mathcal{A}}$). From the proof of Proposition 3.13 (for details see Appendix C), $\Gamma^k_{\mathcal{M}\times\mathcal{A}}$ is (m, ε, κ)-small with $m \leq r$ and $\varepsilon = (\frac{p_{\min}c_{\mathfrak{D}}}{r})^r$, and the result follows from Theorem 3.6 (3.) applied to each $\Gamma^k_{\mathcal{M}\times\mathcal{A}}$ separately. □

5   Conclusions

We have shown that DTA measures over semi-Markov processes are well-defined for almost all runs and assume only finitely many values with positive probability. We also indicated how to approximate DTA measures and the associated probabilities up to an arbitrarily small given precision.

Our approximation algorithm is quite naive and there is a lot of space for further improvement. An interesting open question is whether one can design more efficient algorithms with low complexity in the size of SMP (the size of DTA specifications should stay relatively small in most applications, and hence the (inevitable) exponential blowup in the size of DTA is actually not so problematic).

Another interesting question is whether the results presented in this paper can be extended to more general stochastic models such as generalized semi-Markov processes.

References

[1] R. Alur, C. Courcoubetis, and D.L. Dill. Verifying automata specifications of probabilistic real-time systems. In *Real-Time: Theory in Practice*, volume 600 of *Lecture Notes in Computer Science*, pages 28–44. Springer, 1992.

[2] R. Alur and D. Dill. A theory of timed automata. *Theoretical Computer Science*, 126(2):183–235, 1994. Fundamental Study.

[3] A. Aziz, K. Sanwal, V. Singhal, and R.K. Brayton. Model-checking continuous-time Markov chains. *ACM Transactions on Computational Logic*, 1(1):162–170, 2000.

[4] C. Baier, N. Bertrand, P. Bouyer, T. Brihaye, and M. Größer. Almost-sure model checking of infinite paths in one-clock timed automata. In *Proceedings of LICS 2008*, pages 217–226. IEEE Computer Society Press, 2008.

[5] C. Baier, B.R. Haverkort, H. Hermanns, and J.-P. Katoen. Model-checking algorithms for continuous-time Markov chains. *IEEE Transactions on Software Engineering*, 29(6):524–541, 2003.

[6] V.L. Barbu and N. Limnios. *Semi-Markov Chains and Hidden Semi-Markov Models toward Applications.* Springer, 2008.

[7] N. Bertrand, P. Bouyer, T. Brihaye, and N. Markey. Quantitative model-checking of one-clock timed automata under probabilistic semantics. In *Proceedings of 5th Int. Conf. on Quantitative Evaluation of Systems (QEST'08)*, pages 55–64. IEEE Computer Society Press, 2008.

[8] P. Billingsley. *Probability and Measure*. Wiley, 1995.

[9] D. Bini, G. Latouche, and B. Meini. *Numerical methods for Structured Markov Chains*. Oxford University Press, 2005.

[10] T. Brázdil, J. Krčál, J. Křetínský, A. Kučera, and V. Řehák. Stochastic real-time games with qualitative timed automata objectives. In *Proceedings of CONCUR 2010*, volume 6269 of *Lecture Notes in Computer Science*, pages 207–221. Springer, 2010.

[11] P. Brémaud. *Markov chains: Gibbs fields, Monte Carlo simulation, and queues*. Springer, 1998.

[12] T. Chen, T. Han, J.-P. Katoen, and A. Mereacre. Quantitative model checking of continuous-time Markov chains against timed automata specifications. In *Proceedings of LICS 2009*, pages 309–318. IEEE Computer Society Press, 2009.

[13] E.A. Emerson and E.M. Clark. Using branching time temporal logic to synthesize synchronization skeletons. *Science of Computer Programming*, 2:241–266, 1982.

[14] B.R. Haverkort, H. Hermanns, and J.-P. Katoen. On the use of model checking techniques for quantitative dependability evaluation. In *Proceedings of IEEE Symp. Reliable Distributed Systems*, pages 228–238, 2000.

[15] O. Kallenberg. *Foundations of Modern Probability*. Probability and Its Applications. Springer, 1997.

[16] R. Kress. *Numerical Analysis*. Springer, 1998.

[17] K. Matthes. Zur Theorie der Bedienungsprozesse. *Transactions of the Third Prague Conference on Information Theory, Statistical Decision Functions, Random Processes*, pages 513–528, 1962.

[18] S. Meyn and R.L. Tweedie. *Markov Chains and Stochastic Stability*. Cambridge University Press, 2009.

[19] D.S. Mitrinovic, J. Sándor, and B. Crstici. *Handbook of number theory*. Kluwer, 1996.

[20] J.R. Norris. *Markov Chains*. Cambridge University Press, 1998.

[21] G.O. Roberts and J.S. Rosenthal. General state space Markov chains and MCMC algorithms. *Probability Surveys*, 1:20–71, 2004.

[22] J.S. Rosenthal. *A first look at rigorous probability theory*. World Scientific Publishing, 2006.

[23] S.M. Ross. *Stochastic Processes*. Wiley, 1996.

A   Proof of Corollary 3.2

COROLLARY 3.2. $\mathbf{c}^{\mathcal{A}}$ *is well-defined for almost all runs of* $\mathcal{M}$*. Further, there are pairwise disjoint sets* $\mathcal{R}_1, \ldots, \mathcal{R}_K$ *of runs in* $\mathcal{M}$ *such that* $\mathcal{P}(\mathcal{R}_1 \cup \cdots \cup \mathcal{R}_K) = 1$*, and for every* $1 \leq j \leq K$ *there is a tuple* C_j *such that* $\mathbf{c}^{\mathcal{A}}(\sigma) = C_j$ *for almost all* $\sigma \in \mathcal{R}_j$*.*

Most of the proof has already been presented in Section 3. It remains to prove that for almost all runs σ of $\mathcal{R}_j$ we have

$$\frac{\sum_{s \in S} E_s \cdot D_{j,(s,q)}}{\sum_{p \in Q} \sum_{s \in S} E_s \cdot D_{j,(s,p)}} = \mathbf{c}_q^{\mathcal{A}}(\sigma) \tag{3}$$

To simplify our notation we write $D_{s,p}$ and $1_{s,p}^i$ instead of $D_{j,(s,p)}$ and $1_{(s,p)}^i$, respectively. If $D_q = 0$ then clearly both sides of Equation (3) are 0. Assume that $D_q > 0$.

We prove that for almost all runs σ of $\mathcal{R}_j$,

$$\sum_{s \in S} E_s \cdot D_{s,q} = \lim_{n \to \infty} \frac{\sum_{i=1}^n T^i(\sigma) \cdot 1_q^i(\sigma)}{n} \tag{4}$$

$$\sum_{p \in Q} \sum_{s \in S} E_s \cdot D_{s,p} = \lim_{n \to \infty} \frac{\sum_{i=1}^n T^i(\sigma)}{n} \tag{5}$$

which proves Equation (3) because

$$\lim_{n \to \infty} \frac{\sum_{i=1}^n T^i(\sigma) \cdot 1_q^i(\sigma)}{\sum_{i=1}^n T^i(\sigma)} = \mathbf{c}_q^{\mathcal{A}}(\sigma) \, .$$

By the strong law of large numbers, for almost all runs σ of $\mathcal{R}$ we have

$$E_s = \lim_{n \to \infty} \frac{\sum_{i=1}^n T^i(\sigma) \cdot 1_{s,p}^i(\sigma)}{\sum_{i=1}^n 1_{s,p}^i(\sigma)} \tag{6}$$

for all $s \in S$ and $p \in Q$ satisfying $D_{s,p} > 0$ (note that waiting times in s do not depend on p). Let σ be a run of $\mathcal{R}$ which satisfies Equation (6) for all $s \in S$ and $p \in Q$ where $D_{s,p} > 0$ and such that $\mathbf{d}^{\mathcal{A} \times S}$ is well-defined for σ.

For every $p \in Q$ we have that $\sum_{s \in S} E_s \cdot D_{s,p}$ is equal to

$$
\sum_{s \in S} \lim_{n \to \infty} \frac{\sum_{i=1}^{n} T^i(\sigma) \cdot 1_{s,p}^i(\sigma)}{\sum_{i=1}^{n} 1_{s,p}^i(\sigma)} \cdot D_{s,p} =
$$

$$
= \sum_{s \in S} \lim_{n \to \infty} \frac{\sum_{i=1}^{n} T^i(\sigma) \cdot 1_{s,p}^i(\sigma)}{\sum_{i=1}^{n} 1_{s,p}^i(\sigma)} \cdot \frac{\sum_{i=1}^{n} 1_{s,p}^i(\sigma)}{n}
$$

$$
= \sum_{s \in S} \lim_{n \to \infty} \frac{\sum_{i=1}^{n} T^i(\sigma) \cdot 1_{s,p}^i(\sigma)}{n}
$$

$$
= \lim_{n \to \infty} \frac{\sum_{i=1}^{n} T^i(\sigma) \cdot \sum_{s \in S} 1_{s,p}^i(\sigma)}{n}
$$

$$
= \lim_{n \to \infty} \frac{\sum_{i=1}^{n} T^i(\sigma) \cdot 1_{p}^i(\sigma)}{n}
$$

which proves Equation (4). Also $\sum_{p \in Q} \sum_{s \in S} E_s \cdot D_{s,p}$ is equal to

$$
\sum_{p \in Q} \lim_{n \to \infty} \frac{\sum_{i=1}^{n} T^i(\sigma) \cdot 1_{p}^i(\sigma)}{n} = \lim_{n \to \infty} \frac{\sum_{i=1}^{n} T^i(\sigma) \cdot \sum_{p \in Q} 1_{p}^i(\sigma)}{n}
$$

$$
= \lim_{n \to \infty} \frac{\sum_{i=1}^{n} T^i(\sigma)}{n}
$$

which proves Equation (5) and finishes the proof. $\qquad \square$

B   Proofs of Section 3.2

B.1   Proof of Lemma 3.7

LEMMA 3.7. *Let $A \in \mathcal{G}_{\mathcal{M} \times \mathcal{A}}$. Then $P_{\mathcal{M} \times \mathcal{A}}(\cdot, A)$ is a measurable function, i.e., $\mathcal{M} \times \mathcal{A}$ is a GSSMC.*

Proof. To prove this lemma, it is sufficient to show that $P_{\mathcal{M} \times \mathcal{A}}(\cdot, A)$ is a measurable function from $\Gamma_{\mathcal{M} \times \mathcal{A}}$ to $[0, 1]$ where A ranges (only) over the generators of $\mathcal{G}_{\mathcal{M} \times \mathcal{A}}$, i.e. $A = \{s'\} \times \{q'\} \times \mathbf{I}$ where $s' \in S$, $q' \in Q$, and $\mathbf{I} = \prod_{x \in \mathcal{X}} I_x$ such that I_x is an interval for each $x \in \mathcal{X}$ (see, e.g., [15, Lemma 1.37]).

As the sets S and Q are finite, our goal is to show that a function $P_{\mathcal{M} \times \mathcal{A}}((s, q, \cdot), \{s'\} \times \{q'\} \times \mathbf{I})$ is measurable for $s, s' \in S$, $q, q' \in Q$, and a product of intervals $\mathbf{I}$.

The rest of the proof is based on the fact that a real valued function is measurable, if it is piecewise continuous. Hence, we finish the proof showing that the function $P_{\mathcal{M} \times \mathcal{A}}((s, q, \cdot), \{s'\} \times \{q'\} \times \mathbf{I})$ is piecewise continuous when we fix valuation of all clocks but one. Formally, we fix a valuation ν and a clock

25

x and show that the following function of a parameter $u \in \mathbb{R}_{\geq 0}$ is piecewise continuous.

$$P_{\mathcal{M} \times \mathcal{A}}((s, q, \nu[x := u]), \{s'\} \times \{q'\} \times \mathbf{I}) = \delta_{q'\bar{q}} \cdot \mathbf{P}(s)(s') \cdot \int_0^\infty f(t) \cdot 1_{\mathbf{I}}(\bar{\nu} + t) \, dt$$

where

- $\nu[x := u]$ is the valuation ν where the value of the clock x is set to u;

- δ is the Kronecker delta, i.e., $\delta_{q'\bar{q}} = 1$ if $q' = \bar{q}$, and 0, otherwise;

- $f = \mathbf{D}(s, s')$ is the delay density function for (s, s');

- $(\bar{q}, \bar{\nu})$ is the timed automaton successor of the state $(s, q, \nu[x := u])$, i.e., $\mathcal{A}((s, q, \nu[x := u])) = (\bar{q}, \bar{\nu})$;

- $1_{\mathbf{I}}$ is the indicator function of the set $\mathbf{I}$, i.e., $1_{\mathbf{I}}(\nu') = 1$ if $\nu' \in \mathbf{I}$, and 0, otherwise.

The function $\mathbf{P}(s)(s')$ is constant (recall that s and s' are fixed). Due to the standard region construction for $\mathcal{A}$, it holds that $\bar{q}$ is piecewise constant and $\bar{\nu}$ is piecewise continuous with respect to u.

Let u be in one of the finitely many intervals where $\delta_{q'\bar{q}}$ is constant and the valuation $\bar{\nu}$ changes continuously, i.e. the automaton $\mathcal{A}$ uses the same transition for all u of this interval. As $1_{\mathbf{I}}$ is the indicator function and $\mathbf{I}$ is a product of intervals, it holds that

$$\delta_{q'\bar{q}} \cdot \mathbf{P}(s)(s') \cdot \int_0^\infty f(t) \cdot 1_{\mathbf{I}}(\bar{\nu} + t) \, dt = \delta_{q'\bar{q}} \cdot \mathbf{P}(s)(s') \cdot \int_{\mathbf{a}(u)}^{\mathbf{b}(u)} f(t) \, dt$$

where $\mathbf{a}(u)$ and $\mathbf{b}(u)$ are continuous functions of u and so $\int_{\mathbf{a}(u)}^{\mathbf{b}(u)} f(t) \, dt$ is also a continuous function of u (recall that $\int_0^\infty f(t) \, dt = 1$). Therefore, the function $P_{\mathcal{M} \times \mathcal{A}}((s, q, \nu[x := u]), \{s'\} \times \{q'\} \times \mathbf{I})$ is a piecewise continuous function of u and $P_{\mathcal{M} \times \mathcal{A}}(\cdot, A)$ is a measurable function, i.e., $\mathcal{M} \times \mathcal{A}$ is a GSSMC. $\quad\square$

B.2 Proof of Lemma 3.8

LEMMA 3.8. *There is a measurable one-to-one mapping ξ from the set of runs of $\mathcal{M}$ to the set of runs of $\mathcal{M} \times \mathcal{A}$ such that*

- *ξ preserves measure, i.e., for every measurable set X of runs of $\mathcal{M}$ we have that $\xi(X)$ is also measurable and $\mathcal{P}_{\mathcal{M}}(X) = \mathcal{P}_{\mathcal{M} \times \mathcal{A}}(\xi(X))$;*

- *ξ preserves $\mathbf{d}$, i.e., for every run σ of $\mathcal{M}$ and every $q \in Q$ we have that $\mathbf{d}_q(\sigma)$ is well-defined iff $\mathbf{d}_q(\xi(\sigma))$ is well-defined, and $\mathbf{d}_q(\sigma) = \mathbf{d}_q(\xi(\sigma))$.*

Proof. First, we define the function ξ. We use auxiliary functions $\xi_0 : S \to \Gamma_{\mathcal{M} \times \mathcal{A}}$ that maps the initial states and a function $\xi_\to : \Gamma_{\mathcal{M} \times \mathcal{A}} \times (S \times \mathbb{R}_{>0} \times S) \to \Gamma_{\mathcal{M} \times \mathcal{A}}$ that maps transitions. First, we set $\xi_0(s) = (s, q_0, \mathbf{0})$ where q_0 is the initial location and $\mathbf{0}$ is the zero vector. Next, let $s, s' \in S$, $t \in \mathbb{R}_{>0}$ and $z = (s'', q, \nu) \in \Gamma_{\mathcal{M} \times \mathcal{A}}$. We define $\xi_\to(z, (s, t, s')) = (s', q', \nu' + t)$ such that $\mathcal{A}(z) = (q', \nu')$.

For a run $\sigma = s_0 t_0 s_1 t_1 \cdots$, we use these two functions to set $\xi(\sigma) = z_0 z_1 z_2 \cdots$ such that $\xi_0(s_0) = z_0$ and for each $i \in \mathbb{N}_0$ it holds that $\xi_\to(z_i, (s_i, t_i, s_{i+1})) = z_{i+1}$. We need to show the following claims about the function ξ.

Claim B.1. *Let σ be a run of $\mathcal{M}$. We have for any $q \in Q$ that $\mathbf{d}_q(\sigma)$ is well-defined if and only if $\mathbf{d}_q(\xi(\sigma))$ is well-defined, and $\mathbf{d}_q(\sigma) = \mathbf{d}_q(\xi(\sigma))$.*

Let $\sigma = s_0 t_0 s_1 t_1 s_2 t_2 \cdots$ be a run of $\mathcal{M}$. Let us fix a location $q \in Q$. Recall the Figure 2. The run of $\mathcal{A}$ over σ is a sequence

$$\mathcal{A}(\sigma) = (q_0, \nu_0) s_0 (q_1, \bar{\nu}_0) t_0 (q_1, \nu_1) s_1 (q_2, \bar{\nu}_1) t_1 (q_2, \nu_2) s_2 \cdots .$$

The corresponding run of the product is

$$\xi(\sigma) = (s_0, q_0, \nu_0)(s_1, q_1, \nu_1)(s_2, q_2, \nu_2) \cdots .$$

The values $\mathbf{d}_q(\sigma)$ and $\mathbf{d}_q(\xi(\sigma))$ are limit superior of partial sums of ratio of q in a sequence of locations. For $\mathbf{d}_q(\sigma)$ the sequence is $Q^1(\sigma), Q^2(\sigma), Q^3(\sigma), \ldots = q_1, q_2, q_3, \ldots$ (recall that $Q^i(\sigma)$ is the location entered after reading the finite prefix $s_0 t_0 \cdots s_i$) and for $\mathbf{d}_q(\xi(\sigma))$ the sequence is also $q_1, q_2, q_3, \ldots$. Hence, we get that $\mathbf{d}_q(\sigma)$ is well-defined iff $\mathbf{d}_q(\xi(\sigma))$ is well-defined and $\mathbf{d}_q(\sigma) = \mathbf{d}_q(\xi(\sigma))$.

Claim B.2. *For any measurable set X of runs of $\mathcal{M}$, the set $\xi(X)$ is measurable.*

Recall that by $\mathcal{R}(B)$ we denote a cylinder of runs that follow the given template B. Let X be a set of runs such that $X = \mathcal{R}(B)$ for some template $B = s_0 I_0 \cdots s_n I_n$, i.e. X is from the generator set. We can cover the image of X by cylinders composed of basic hybercubes. By decreasing the edge length of the hypercubes to the limit, we then get a set that equals the image of X. For $k \in \mathbb{N}$ and $\mathbf{v} \in (\mathbb{N}_0)^{|\mathcal{X}|}$ we denote by $C_\mathbf{v}^k$ a set of valuations $\prod_{x \in \mathcal{X}} [\mathbf{v}(x)/k, (\mathbf{v}(x)+1)/k]$. The set of all cylinder templates composed of basic hypercubes of precision k is

$$U_k = \{A_0 \cdots A_n \mid A_i = \{s_i\} \times \{q_i\} \times C_{\mathbf{v}_i}^k, s_i \in S, q_i \in Q, \mathbf{v}_i \in (\mathbb{N}_0)^{|X|}\}$$

A run $\sigma = z_0 z_1 \cdots$ of $\mathcal{M} \times \mathcal{A}$ is in $\mathcal{R}(A_0 \ldots A_n)$ if for each $0 \leq i \leq n$ we have $z_i \in A_i$. It is easy to show that

$$\xi(X) = \bigcap_{k \in \mathbb{N}} \bigcup \{\mathcal{R}(C) \mid C \in U_k, \mathcal{R}(C) \cap \xi(X) \neq \emptyset\}$$

hence, $\xi(X)$ is a measurable set. By standard arguments we get the result for any measurable X.

Claim B.3. *For any measurable set X of runs of $\mathcal{M} \times \mathcal{A}$, the set $\xi^{-1}(X)$ is measurable.*

The arguments are similar as in the previous claim. Let Y be a set of runs such that $Y = \mathcal{R}(C)$ for some template

$$C = \{s_0\} \times \{q_0\} \times \prod_{x \in \mathcal{X}} I_{x,0} \quad \cdots \quad \{s_n\} \times \{q_n\} \times \prod_{x \in \mathcal{X}} I_{x,n},$$

i.e. Y is from the generator set. By I_i^k we denote an interval $[i/k, (i+1)/k]$. The set of all cylinder templates in $\mathcal{M}$ composed of basic lines of precision k is

$$T_k = \{s_0 I_{i_0}^k \cdots s_n I_{i_n}^k \mid s_i \in S, i_n \in \mathbb{N}_0\}$$

Again, it is easy to show that

$$\xi^{-1}(Y) = \bigcap_{k \in \mathbb{N}} \bigcup \{\mathcal{R}(B) \mid B \in T_k, \mathcal{R}(B) \cap \xi^{-1}(Y) \neq \emptyset\}$$

hence, $\xi^{-1}(Y)$ is a measurable set. Again, by standard arguments we get the result for any measurable X.

Claim B.4. *For any measurable set X of runs of $\mathcal{M}$, we have $\mathcal{P}_{\mathcal{M}}(X) = \mathcal{P}_{\mathcal{M} \times \mathcal{A}}(\xi(X))$.*

We define a new measure $\mathcal{P}'_{\mathcal{M}}$ over runs of $\mathcal{M}$ by

$$\mathcal{P}'_{\mathcal{M}}(X) = \mathcal{P}_{\mathcal{M} \times \mathcal{A}}(\xi(X))$$

for any measurable set of runs X. First we need to show that $\mathcal{P}'_{\mathcal{M}}$ is a probability measure, i.e. $\mathcal{P}'_{\mathcal{M}}(\emptyset) = 0$, $\mathcal{P}'_{\mathcal{M}}(\mathcal{R}_{\mathcal{M}}) = 1$, and for any collection of pairwise disjoint sets $X_1, \ldots, X_n$ we have $\mathcal{P}'_{\mathcal{M}}(\bigcup_{i=1}^n X_i) = \sum_{i=1}^n \mathcal{P}'_{\mathcal{M}}(X_i)$. The first and the second statement follows directly from the definition of ξ, the third statement follows from the fact that $\mathcal{P}_{\mathcal{M} \times \mathcal{A}}$ satisfies this property and that ξ-image of disjoint sets are disjoint sets of runs which can be easily checked.

Let $B = s_0 I_0 \cdots s_n I_n$ be a cylinder template. We show

$$\mathcal{P}'_{\mathcal{M}}(\mathcal{R}(B)) = \mathcal{P}_{\mathcal{M}}(\mathcal{R}(B)).$$

We obtain $\mathcal{P}'_{\mathcal{M}} = \mathcal{P}_{\mathcal{M}}$ by the extension theorem because $\mathcal{P}'_{\mathcal{M}}$ and $\mathcal{P}_{\mathcal{M}}$ coincide on the generators. From the definition of $\mathcal{P}'_{\mathcal{M}}$ we get the claim. From the definition of semi-Markov process, we have

$$\mathcal{P}_{\mathcal{M}}(\mathcal{R}(B)) = \alpha_0(s_0) \cdot \prod_{i=0}^n \mathbf{P}(s_i)(s_{i+1}) \cdot \int_{t_i \in I_i} f_i(t_i) \, dt_i$$

where $f_i = \mathbf{D}(s_i, s_{i+1})$ is the density of the transition from s_i to s_{i+1}. Now we turn our attention to the product. For the fixed template we define a set $N_0 = \{(s_0, q_0, \mathbf{0})\}$ and a sequence of functions $N_1, \ldots, N_n$ such that

$$N_{i+1}(z_i) = \{\xi_{\rightarrow}(z_i, (s_i, t_i, s_{i+1})) \mid t_i \in I_i\}.$$

For an interval $I = [a, b]$ and valuation ν we define $C(I, \nu)$ to be a hypercube of edge length $b - a$ starting at $\nu + a$, i.e. $C(I, \nu) = \prod_{x \in \mathcal{X}}[\nu(x) + a, \nu(x) + b]$. Now for each $i \in \mathbb{N}_0$ the conditional density

$$\mathcal{P}_{\mathcal{M} \times \mathcal{A}}(\Phi_{i+1} \in N_{i+1}(\Phi_i) \mid \Phi_i)$$
$$= \mathcal{P}_{\mathcal{M} \times \mathcal{A}}(\Phi_{i+1} \in \bar{N}_{i+1}(\Phi_i) \mid \Phi_i)$$

where $\bar{N}_{i+1}(z) = \bigcup_{q_{i+1} \in Q}\{s_{i+1}\} \times \{q_{i+1}\} \times C(I_i, \bar{\nu})$ where $\mathcal{A}(z) = (\bar{q}, \bar{\nu})$. The equality holds because $N_{i+1}(\Phi_i) \subseteq \bar{N}_{i+1}(\Phi_i)$ and $\mathcal{P}_{\mathcal{M} \times \mathcal{A}}(\Phi_{i+1} \in (\bar{N}_{i+1}(\Phi_i) \setminus N_{i+1}(\Phi_i)) \mid \Phi_i) = 0$. Indeed, the probability of hitting anything else but the diagonal of the hypercube $\bar{N}_{i+1}(\Phi_i)$ is clearly 0. Because $\bar{N}_{i+1}(z)$ is for each z a union of basic cylinders for that we have explicit definition of the transition kernel, we have

$$= \sum_{q \in Q} \delta_{q\bar{q}} \cdot \mathbf{P}(s_i)(s_{i+1}) \cdot \int_0^\infty f_i(t) \cdot 1_{C(I_i, \bar{\nu})}(\bar{\nu} + t) dt$$

where δ is the Kronecker delta, i.e., $\delta_{q\bar{q}} = 1$ if $q = \bar{q}$, and 0, otherwise. Notice that here $\bar{q}$ and $\bar{\nu}$ are random variables such that $\mathcal{A}(\Phi_i) = (\bar{q}, \bar{\nu})$. The rest of the formula after $\delta_{q\bar{q}}$ does not depend on $\bar{q}$, we can write

$$= \mathbf{P}(s_i)(s_{i+1}) \cdot \int_0^\infty f_i(t) \cdot 1_{C(I_i, \bar{\nu})}(\bar{\nu} + t) dt.$$

Furthermore, hitting the hypercube $C(I_i, \bar{\nu})$ equals to waiting for a time from I_i

$$= \mathbf{P}(s_i)(s_{i+1}) \cdot \int_{I_i} f_i(t) dt.$$

Hence, the conditioned probability is a constant random variable that does not depend on Φ_i. Finally,

$$
\begin{aligned}
\mathcal{P}'_{\mathcal{M}}(\mathcal{R}(B)) &= \mathcal{P}_{\mathcal{M} \times \mathcal{A}}(\xi(\mathcal{R}(B))) \\
&= \mathcal{P}_{\mathcal{M} \times \mathcal{A}}(\Phi_0 \in N_0, \Phi_1 \in N_1(\Phi_0), \ldots, \Phi_n \in N_n(\Phi_{n-1})) \\
&= \mathcal{P}_{\mathcal{M} \times \mathcal{A}}(\Phi_n \in N_n(\Phi_{n-1}) \mid \Phi_0, \ldots, \Phi_{n-1}) \cdot \cdots \\
&\quad \cdot \mathcal{P}_{\mathcal{M} \times \mathcal{A}}(\Phi_1 \in N_n(\Phi_0) \mid \Phi_0) \cdot \mathcal{P}_{\mathcal{M} \times \mathcal{A}}(\Phi_0 \in N_0) \\
&= \prod_{i=0}^n \left(\mathbf{P}(s_i)(s_{i+1}) \cdot \int_{t_i \in I_i} f_i(t_i) dt_i \right) \cdot \alpha_0(s_0) \\
&= \mathcal{P}_{\mathcal{M}}(\mathcal{R}(B))
\end{aligned}
$$

which concludes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

C Proofs of Section 3.3

C.1 Proofs of Proposition 3.11

PROPOSITION 3.11. *Assume that $\mathcal{G}_{\mathcal{M} \times \mathcal{A}}$ is strongly connected and aperiodic. Then there exists a region R and $S \subseteq R$, $n \in \mathbb{N}$, $b > 0$, and a probability measure μ such that $\mu(S) = 1$ and for every measurable $T \subseteq S$ and $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}$ we have $P_{\mathcal{M} \times \mathcal{A}}^n(z, T) > b \cdot \mu(T)$. In other words, the set $\Gamma_{\mathcal{M} \times \mathcal{A}}$ of all states of the GSSMC $\mathcal{M} \times \mathcal{A}$ is (n, b, μ)-small.*

Proof. Follows easily from Proposition 3.13 by considering the period p equal to 1. $\qquad\qquad\square$

C.2 Proofs of Proposition 3.13

PROPOSITION 3.13. *Assume that $\mathcal{G}_{\mathcal{M} \times \mathcal{A}}$ is strongly connected and has a period p. For every $k \in \{0, \ldots, p-1\}$ there exists a set of states S_k, $n_k \in \mathbb{N}$, $b_k > 0$, and a probability measure κ_k such that $\kappa_k(S_k) = 1$ and for every measurable $T \subseteq S_k$ and $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}^k$ we have $P_{\mathcal{M} \times \mathcal{A}}^{n_k \cdot p}(z, T) > b_k \cdot \kappa_k(T)$. In other words, Φ_k is (n_k, b_k, κ_k)-small.*

In the following text we formulate the definitions and lemmata needed to prove the proposition. The actual proofs of the lemmata are in next subsections (grouped by proof techniques).

Let us fix a $k \in \{0, \ldots, p-1\}$. We show that there is a state $z^* \in \Gamma_{\mathcal{M} \times \mathcal{A}}^k$ such that for each starting state $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}^k$ there is a path $z \cdots z^*$ of length $n_k \cdot p$ that is δ-*wide*. For a fixed $\delta > 0$, it means that the waiting time of any transition in the path can be changed by $\pm \delta$ without ending up in a different region in the end. Precise definition follows.

Definition C.1. *Let $z = (s, q, \nu)$ and $z' = (s', q', \nu')$ be two states. For a waiting time $t \in \mathbb{R}_{>0}$ we set $z \xrightarrow{t} z'$ if $\mathcal{A}(z) = (q', \bar{\nu})$ and $\nu' = \bar{\nu} + t$. We set $z \to z'$ if $z \xrightarrow{t} z'$ for some $t \in \mathbb{R}_{>0}$ and call it a feasible transition.*

For $\delta > 0$, we say that a feasible transition $z \to z'$ is δ-wide if for every $x \in \mathcal{X}$ relevant for ν' we have $\mathit{frac}(\nu_i(x)) \in [\delta, 1 - \delta]$.

Let $z_1 \cdots z_n$ be a path. It is feasible if for each $1 \le i < n$ we have that $z_i \to z_{i+1}$. It is δ-wide if for each $1 \le i < n$ we have that $z_i \to z_{i+1}$ is a δ-wide transition.

By next lemma, we reduce the proof of Proposition 3.13 to finding δ-wide paths from any z to the fixed z^*.

First, we recall the following notation that is necessary for analyzing the computational complexity. Let $p_{\min}$ denote the smallest probability in $\mathcal{M}$. Further, let us denote by $\mathfrak{D}(\mathcal{M})$ the set of delay densities used in $\mathcal{M}$, i.e. $\mathfrak{D}(\mathcal{M}) = \{ \mathbf{D}(s, s') \mid s, s' \in S \}$. From our assumptions imposed on delay densities we obtain the following uniform bound $c_{\mathfrak{D}} > 0$ on delay densities of $\mathfrak{D}(\mathcal{M})$. For every $f \in \mathfrak{D}(\mathcal{M})$ and for all $x \in [0, B_{\max}]$, either $f(x) > c_{\mathfrak{D}}$ or $f(x) = 0$, and moreover, $\int_{B_{\max}}^\infty f(x) dx > c$ or equals 0.

Lemma C.2. *For every $\delta > 0$ and $n > 1$ there is a probabilistic measure κ and $b > 0$ such that the following holds. For every δ-wide path $\sigma = z_0 z_1 \cdots z_n$, there*

is a κ-measurable set of states Z with $\kappa(Z) = 1$ such that $z_n \in Z$ and for any measurable subset $Y \subseteq Z$ it holds $P^n_{\mathcal{M} \times \mathcal{A}}(z_1, Y) \geq b \cdot \kappa(Y)$.

Moreover, we can set $b = (p_{\min} \cdot c_{\mathfrak{D}} \cdot \delta/n)^n / \sqrt{|\mathcal{X}|}$.

Now it remains to find a state z^* and a δ-wide path to z^* for any z. Such path is composed of five parts, each having a fixed length. The target state z^* is then the first state where all these paths from all starting states z meet together.

In the first part, we move to a δ'-*separated* state for some $\delta' > 0$.

Definition C.3. *Let $\delta > 0$. We say that a set $X \subseteq \mathbb{R}_{\geq 0}$ is δ-separated if for every $x, y \in X$ either $frac(x) = frac(y)$ or $|frac(x) - frac(y)| > \delta$.*

Further, we say that $(s, q, \nu) \in \Gamma_{\mathcal{M} \times \mathcal{A}}$ is δ-separated if the set

$$\{0\} \cup \{\nu(x) \mid x \in \mathcal{X}, x \text{ is relevant for } \nu\}$$

is δ-separated.

Now we can formulate the first part of the path precisely.

Lemma C.4. *There is $\delta > 0$ and $n \in \mathbb{N}$ such that for any $z_1 \in \Gamma_{\mathcal{M} \times \mathcal{A}}$ there is a δ-wide path $z_1 \cdots z_n$ such that z_n is δ-separated.*
Moreover, we can set $n = B_{\max} \cdot |\mathcal{X}|$ and $\delta = 1/(2(|\mathcal{X}| + 2))$.

At the beginning of the second part, we are in a δ-separated state z_1 in some region $R \in V_{k'}$ for some $k' \in \{0, \ldots, p-1\}$. For the given k', we fix a region $R_1 \in V_{k'}$. Due to strong connectedness, reaching R_1 is possible from any state in $V_{k'}$ in a fixed sufficiently large number of steps n'. By a path of length n' that is (δ/n')-wide, we reach a (δ/n')-separated state in R_1. The separation and wideness decreases with length because the fractional values of relevant clock may be forced to get closer and closer to each other by resets on the path to R_1. The reason for the first part of the path was only to bound the wideness of the second part.

Lemma C.5. *Let the region graph $G_{\mathcal{M} \times \mathcal{A}}$ be strongly connected and let p be the period of $G_{\mathcal{M} \times \mathcal{A}}$. Let $k \in \{0, \ldots, p-1\}$, $\delta > 0$ and $R \in V_k$ be a region. Then there is $n \in \mathbb{N}$ such that for every δ-separated $z_1 \in \Gamma^k_{\mathcal{M} \times \mathcal{A}}$ there is a (δ/n)-wide path $z_1 \cdots z_n$ such that z_n is (δ/n)-separated and $z_n \in R$.*
Moreover, we can set $n = \lfloor |V|^{4 \ln |V| - 1}/6 \rfloor \cdot p$.

In the third part, we squeeze the fractional values of all relevant clocks close to 0. We go through a fixed region path $R_1 \cdots R_k$ such that in each step we shift the time by an integral value minus a small constant c. This way the reset clocks are fractionally placed to 0 and the other clocks decrease their fractional values only by the small constant c. Since we go through a fixed region path, we have a fixed sequence of sets of clocks $\mathcal{X}_1, \ldots, \mathcal{X}_{k-1}$ reset in respective steps. Hence, the fractional values of clocks reset during this path have fixed relative distances. For any starting state z_1' we reach a state z_k' that *almost equals* a fixed "reference" state $z_k \in R_k$.

Definition C.6. *Let $z, z' \in \Gamma_{\mathcal{M} \times \mathcal{A}}$. We say that state z almost equals state z' if $z \sim z'$ and each clock relevant in z has the same value in z and z'.*

Notice that clocks not relevant in z_k may still have different values.

Lemma C.7. *Let R be a region. For each $\delta > 0$ there is $\delta' > 0$, $n \in \mathbb{N}$ and $z' \in \Gamma_{\mathcal{M} \times \mathcal{A}}$ such that for every δ-separated $z_1 \in R$ there is a δ'-wide path $z_1 \cdots z_n$ such that z_n and $\bar{z}$ almost equal.*
Moreover, we can set $n = B_{\max} + 1$ and $\delta' = \delta/(B_{\max} + 2)$.

In the fourth part, we somewhat repeat the first part and prepare for the fifth part. We reach a δ-separated state that is almost equal to a fixed state $z_l \in R_l$. Again, we do it by a δ-wide path.

Lemma C.8. *Let z be a state. There is a $\delta > 0$, $n \in \mathbb{N}_0$, and z' such that for any state z_1 almost equal to z there is a δ-wide path $z_1 \cdots z_n$ such that z_n is δ-separated and z_n almost equals z'.*
Moreover, we can set $n = B_{\max} \cdot |\mathcal{X}|$ and $\delta = 1/(2(|\mathcal{X}| + 2))$.

In the fifth part, we go through a fixed region path $R_l \cdots R_{l+m}$ such that each clock not relevant in R_l is reset during this path and hence we reach a fixed $z^* \in R_{l+m}$. Such path exists from the assumption that it is possible to reset every clock. The (arbitrary) values of clocks not relevant in R_l do not influence the behavior of the timed automaton before their reset and we indeed follow a fixed region path. Furthermore, we can stretch the path to arbitrary length so that the length of the whole path is a multiple of the period p. Again, the fifth part of the path is δ/n''-wide where n'' is the number of steps.

Lemma C.9. *Let the region graph $G_{\mathcal{M} \times \mathcal{A}}$ be strongly connected. Let $\delta > 0$. Let z be a δ-separated state. Then there is $n \in \mathbb{N}$, such that for any $n' \geq n$ there is a state z^* such that the following holds. For any state z_1 almost equal to z there is a (δ/n)-wide path $z_1 \cdots z_{n'}$ such that $z_{n'} = z^*$.*
Moreover, we can set $n = |V| \cdot |\mathcal{X}|$.

Now we can finally prove the main proposition.

of Proposition 3.13. We fix $k \in \{0, \ldots, p-1\}$. By Lemmata C.4, C.5, C.7, C.8, and C.9 we get for any state $z_1 \in \Gamma_{\mathcal{M} \times \mathcal{A}}^k$ a δ-wide path $z_1 \cdots z_x$ of length $x = n_k \cdot p$ such that

$$x = B_{\max} \cdot |\mathcal{X}| + M + (B_{\max} + 1) + B_{\max} \cdot |\mathcal{X}| + |V| \cdot |\mathcal{X}| + c \leq 2 \cdot M$$

$$\delta = \frac{1}{((B_{\max} + 2) \cdot 2(|\mathcal{X}| + 2) \cdot M} \geq \frac{1}{4 \cdot B_{\max} \cdot |\mathcal{X}| \cdot M}$$

where $M = \lfloor (|V|^{4 \ln |V| - 1})/6 \rfloor \cdot p$ and $c < p$ is the constant such that x is a multiple of p (we stretch the path by c in the fifth part). Therefore, by Lemma C.2, Φ_k is (n_k, b_k, κ_k)-small for $n_k \leq \lfloor |V|^{4 \ln |V| - 1} \rfloor \cdot p \leq \lfloor |V|^{4 \ln |V|} \rfloor =: r$ and

$$b_k = (p_{\min} \cdot c_{\mathfrak{D}} \cdot \delta/x)^x / \sqrt{|\mathcal{X}|}$$

$$\geq \left(\frac{p_{\min} \cdot c_{\mathfrak{D}}}{8 \cdot B_{\max} \cdot |\mathcal{X}| \cdot M^2} \right)^{2M} \cdot \frac{1}{\sqrt{|\mathcal{X}|}}$$

$$\geq \left(\frac{p_{\min} \cdot c_{\mathfrak{D}}}{M^3} \right)^{2M} \geq \left(\frac{p_{\min} \cdot c_{\mathfrak{D}}}{r} \right)^r$$

Notice that in the calculation above we ignore the cases of trivially small region graphs with less than 3 vertices. □

From the proof we directly get the following bound on constants.

Corollary C.10. *Assume that $\mathcal{G}_{\mathcal{M}\times\mathcal{A}}$ is strongly connected and has a period p. For every $k \in \{0,\ldots,p-1\}$ we have Φ_k is (n,b,κ)-small for some $n \leq r$ divisible by p, and $b = (p_{\min} \cdot c_{\mathfrak{D}}/r)^r$, where $r = \lfloor |V|^{4\ln|V|} \rfloor$.*

C.2.1 Proof of Lemma C.2

LEMMA C.2. *For every $\delta > 0$ and $n > 1$ there is a probabilistic measure κ and $b > 0$ such that the following holds. For every δ-wide path $\sigma = z_0 z_1 \cdots z_n$, there is a κ-measurable set of states Z with $\kappa(Z) = 1$ such that $z_n \in Z$ and for any measurable subset $Y \subseteq Z$ it holds $P^n_{\mathcal{M}\times\mathcal{A}}(z_1, Y) \geq b \cdot \kappa(Y)$.*
Moreover, we can set $b = (p_{\min} \cdot c_{\mathfrak{D}} \cdot \delta/n)^n / \sqrt{|\mathcal{X}|}$.

Proof. Recall that we assume that all delays' densities are bounded by some $c_{\mathfrak{D}} > 0$ in the following sense. For every $d \in \mathfrak{D}$ and for all $x \in [0, B]$, $d(x) > c_{\mathfrak{D}}$ or equals 0. Similarly, $\int_B^\infty d(x)dx > c_{\mathfrak{D}}$ or equals 0.

Let $\sigma = z_0 z_1 \cdots z_n = (s_0, q_0, \nu_0)(s_1, q_1, \nu_1)\cdots(s_n, q_n, \nu_n)$. For $1 \leq i \leq n$, let t_i be the waiting times such that $z_{i-1} \xrightarrow{t_i} z_i$, and let $X_i = \{x \in \mathcal{X} \mid \mathcal{A}(z_{i-1}) = (q, \nu), \nu(x) = 0\}$ be the set of clocks reset right before waiting t_i.

For $\varepsilon > 0$, we define an ε-neighbourhood of σ to be the set of paths of the form $z_0 \xrightarrow{t_1'} (s_1, q_1, \nu_1')\cdots \xrightarrow{t_n'} (s_n, q_n, \nu_n')$ where $t_i' \in (t_i - \varepsilon, t + \varepsilon)$. Due to δ-separation of σ, all paths of its δ/n-neighbourhood are feasible. Considering this δ/n-neighbourhood, the set of all possible ν_n's forms the sought set Z. We may compute this set as follows. We define a mapping $\alpha_\sigma : (-\varepsilon, \varepsilon)^n \to \mathbb{R}^{|\mathcal{X}|}_{\geq 0}$ so that $\alpha_\sigma(\zeta_1,\ldots,\zeta_n) = \nu_n'$ for $t_i' = t_i + \zeta_i$. This can be done by setting $\alpha_\sigma(\zeta_1,\ldots,\zeta_n)(x) = \sum_{r_x}^n (t_i + \zeta_i)$, where the clock x was reset in the r_xth step for the last time in σ, i.e. $r_x = \max\{i \mid x \in X_i\}$. Obviously, α_σ is a restriction of a linear mapping. Therefore, $\alpha_\sigma((-\varepsilon, \varepsilon)^n)$ is an open rhombic hypercube of a dimension $1 \leq d \leq |\mathcal{X}|$. Due to the last summand, it has a positive κ_d-measure. (Here κ_d is the standard Lebesgue measure on the d-dimensional affine space that contains $\alpha_\sigma((-\varepsilon, \varepsilon)^n)$. Equivalently, it is the d-dimensional Hausdorff measure multiplied by the volume of unit d-ball.)

We set $Z := \alpha_\sigma((-\delta/2n, \delta/2n)^n)$. Thus, for every $z \in Z$ there is a $\delta/2$-separated path τ from z_0 to z. We need to construct $b > 0$ such that for all $Y \subseteq Z$, we have $P^n_{\mathcal{M}\times\mathcal{A}}(z_0, Y) \geq b\kappa_d(Y)/\kappa_d(Z) =: b\kappa(Y)$. It is sufficient to prove this for generators of the same topology. We pick the generators as follows. For $z \in Z$ and $\varepsilon < \delta/2n$, we denote $Y(z, \varepsilon) = \alpha_\sigma((-\delta/n, \delta/n)^n) \cap C^{|\mathcal{X}|}_{z,\varepsilon}$, where $C^{|\mathcal{X}|}_{z,\varepsilon}$ is a hypercube with dimension $|\mathcal{X}|$ and size ε centered in z. Clearly, the set of all $Y(z, \varepsilon) \subseteq Z$ form a generator set. We now construct $b > 0$ so that for every such $Y := Y(z, \varepsilon)$ we have $P^n_{\mathcal{M}\times\mathcal{A}}(z_0, Y) \geq b\kappa_d(Y)/\kappa_d(Z)$. To this end, we prove later on that

$$P^n_{\mathcal{M}\times\mathcal{A}}(z_0, Y) \quad \geq \quad (p_{\min}c_{\mathfrak{D}}/n)^n \delta^{n-d} \varepsilon^d \qquad (7)$$

33

Since $\kappa_d(Y) \leq \sqrt{|\mathcal{X}|} \cdot (\varepsilon)^d$ and $\kappa_d(Z) \geq (\delta/n)^d$, we can set $b = (p_{\min}c_{\mathfrak{D}}\delta/n)^n/\sqrt{|\mathcal{X}|}$.

It remains to prove (7). Let $k_1, \ldots, k_d$ be the elements of $\{r_x \mid x \in \mathcal{X}\}$ in the increasing order, and $\ell_1, \ldots, \ell_{n-d}$ the remaining numbers in $\{1, \ldots, n\}$. Note that since α is linear, $\alpha^{-1}(Y)$ is λ_n-measurable (λ_n denotes the standard Lebesgue measure on $\mathbb{R}^n$). Intuitively, if we want to make clock x hit Y, it is sufficient to adjust the waiting time after the last reset of x. Let $Y|X_i$ denote the projection of Y to coordinates in X_i (setting other components to zero). The first equation makes use of the facts that (1) all components of each point of $Y|X_i$ have the same value (because Y is a subset of image of α_σ) and (2) when factoring out all (identical) components but one of each X_i, the image of Y is a d-hypercube (due to the intersection with $C_{z,\varepsilon}^{|\mathcal{X}|}$), so we can use projections independently.

$$
\begin{aligned}
P_{\mathcal{M}\times\mathcal{A}}^n(z_0, Y) \;\geq\;& \int_{\{\mathbf{x}|\alpha_\tau(\mathbf{x})\in Y\}} (p_{\min}c_{\mathfrak{D}})^n d\lambda_n \\
=\;& (p_{\min}c_{\mathfrak{D}})^n \int_{t_{\ell_1}-\delta/2n}^{t_{\ell_1}+\delta/2n} \cdots \int_{t_{\ell_{n-d}}-\delta/2n}^{t_{\ell_{n-d}}+\delta/2n} \\
& \int_{\alpha_\tau(0,\ldots,0,d_{k_d},\ldots,d_n)|X_{k_d}\in Y|X_{k_d}} \cdots \int_{\alpha_\tau(0,\ldots,0,d_{k_1},\ldots,d_n)|X_{k_1}\in Y|X_{k_1}} \\
& d\zeta_{k_1}\cdots d\zeta_{k_d} d\zeta_{\ell_{n-d}}\cdots d\zeta_{\ell_1} \\
=\;& (p_{\min}c_{\mathfrak{D}})^n \underbrace{\int_{-\delta/2n}^{\delta/2n} \cdots \int_{-\delta/2n}^{\delta/2n}}_{n-d} \underbrace{\int_{-\varepsilon/2n}^{\varepsilon/2n} \cdots \int_{-\varepsilon/2n}^{\varepsilon/2n}}_{d} \\
& d\zeta_{k_1}\cdots d\zeta_{k_d} d\zeta_{\ell_{n-d}}\cdots d\zeta_{\ell_1} \\
=\;& (p_{\min}c_{\mathfrak{D}})^n (\delta/n)^{n-d}(\varepsilon/n)^d = (p_{\min}c_{\mathfrak{D}}/n)^n \delta^{n-d}\varepsilon^d \quad \square
\end{aligned}
$$

$\square$

C.2.2   Proofs of Lemmata C.4 and C.8

LEMMA C.4. *There is $\delta > 0$ and $n \in \mathbb{N}$ such that for any $z_1 \in \Gamma_{\mathcal{M}\times\mathcal{A}}$ there is a δ-wide path $z_1\cdots z_n$ such that z_n is δ-separated.*
Moreover, we can set $n = B_{\max}\cdot(|\mathcal{X}|+2)$ and $\delta = 1/(2(|\mathcal{X}|+2))$.

Proof. To simplify the argumentation we introduce a notion of a *r-grid* that marks r distinguished points (called *lines*) on the $[0,1]$ line segment. In the proof we show that we can place fractional values of all relevant clocks on such distinguished points. Let $r \in \mathbb{N}$. We say that a set of clocks $\mathcal{Y} \subseteq \mathcal{X}$ is *on r-grid* in z if for every $x \in \mathcal{Y}$ relevant in z we have $frac(\nu(x)) = n/r$ for some $0 \leq n < r$. For $0 \leq n < r$, we say that the *n-th line of the r-grid is free* in z if there is no relevant clock in the $1/2k$-neighborhood of the n-th line, i.e. for any relevant $x \in \mathcal{X}$ we have $frac(\nu(x)) \notin (n/r - 1/2r, n/r + 1/2r)$.

Let $r = |X| + 2$. We inductively build a $1/2r$-wide path $z_1 \cdots z_n$ where $n = B_{\max} \cdot r$. The set $\emptyset$ is on r-grid in z_1. We show that if a set $\mathcal{Y}_i$ is on r-grid in state z_i, there is a $1/2k$-wide transition to z_{i+1} such that $(\mathcal{Y}_i \cup \mathcal{Z})$ is on r-grid in z_{i+1} where $\mathcal{Z}$ is the set of clocks newly reset in z_i. There are $|X| + 2$ lines on the grid and only $|X|$ clocks. At least two of these lines must be free. Let $j \neq 0$ be such a line. Let t be a waiting time and z_{i+1} a state such that $frac(t) = 1 - j/r$ and $z_i \xrightarrow{t} z_{i+1}$. Such waiting time must be indeed possible because the interval where the density function of any transition is positive has integral bounds. The transition $z_i \xrightarrow{t} z_{i+1}$ is $1/2r$-wide because the line j is free in z_i. Furthermore, the set $(\mathcal{Y}_i \cup \mathcal{Z})$ is on r-grid in z_{i+1} because the fractional value of each clock that was previously on r-grid was changed by a multiple of $1/r$. The newly reset clocks have fractional value $1 - j/r$ which is again a multiple of $1/r$.

Next, we show that $\mathcal{X}$ is on r-grid in z_n. Clocks reset in this path on r-grid in z_n. The remaining clocks are all irrelevant because the path of $B_{\max} \cdot r$ steps takes at least $B_{\max}$ time units. Indeed, each transition in this path takes at least $1/r$ time unit. According to the definition, $\mathcal{X}$ is on r-grid in z_n. Hence, the state z_n is $1/r$-separated because the distance between two adjacent grid lines is $1/r$. By setting $\delta = 1/2r$ we get the result. $\qquad\square$

LEMMA C.8. *Let z be a state. There is a $\delta > 0$, $n \in \mathbb{N}_0$, and z' such that for any state z_1 almost equal to z there is a δ-wide path $z_1 \cdots z_n$ such that z_n is δ-separated and z_n almost equals z'.*
Moreover, we can set $n = B_{\max} \cdot |\mathcal{X}|$ and $\delta = 1/(2(|\mathcal{X}| + 2))$.

Proof. Let us fix a state z_1 almost equal to z. By Lemma C.4 we get a δ-wide path $z_1 \ldots z_n$ such that z_n is δ-separated.

Notice that for a fixed state z, control state s and time t there is a unique location q and valuation ν, hence a unique state $z' = (s, q, \nu)$ such that $z \xrightarrow{t} z'$.

Let $t_1, \ldots, t_{n-1}$ be the waiting times and $s_1, \ldots, s_n$ the control states on the path $z_1, \ldots, z_n$. For any $\bar{z}_1$ almost equal to z_1 we can build using the same waiting times and control states a path $\bar{z}_1 \cdots \bar{z}_n$. It is easy to see that for two almost equal states $z, \bar{z}$ a control state s and a time $t > 0$ the states $z', \bar{z}'$ determined by s and t are also almost equal. Inductively, we get that $\bar{z}_n$ is almost equal to z_n. It also holds that $\bar{z}_n$ is δ-separated since δ-separation is defined only with respect to relevant clocks. $\qquad\square$

C.2.3 Proofs of Lemmata C.5 and C.9

For the proof of Lemma C.5 we need the following result from graph theory.

Lemma C.11. *Let G be a strongly connected and aperiodic oriented graph with $N > 2$ vertices. Then for each $n \geq \lfloor N^{4 \ln N - 1}/6 \rfloor$, there is a path of length precisely n between any two vertices of G.*

Proof. It is a standard result from the theory of Markov chains, see e.g. [22, Lemma 8.3.9], that in every ergodic Markov chain there is n_0 such that between any two states there is a path of any length greater than n_0. In the following, we give a simple bound on n_0.

35

Let u, v be vertices. By aperiodicity, there are C cycles on u of lengths $c_1, \ldots, c_C \leq N$ with $\gcd(c_1, \ldots, c_C) = 1$. Thus by Bézout's identity, there are $m_i \in \mathbb{N}_0$ such that $1 = \sum_{i=1}^{C} m_i c_i$. Hence also $1 = \sum_{i=1}^{C} (m_i + k_i/c_i \cdot \prod_{j=1}^{C} c_j) c_i$ for any $k_1 + \cdots + k_C = 0$. Therefore, $1 = \sum_{i=1}^{j} n_i c_i$ with some $0 > n_i > -1/c_i \cdot \prod_{j=1}^{C} c_j$ for $i < C$ and $n_C > 0$. By [11, Theorem A.1.1], n_0 can be chosen $N + P(P-1)$, where $P = \sum_{i=1}^{C-1} |n_i| c_i$, i.e. the absolute value of the negative part of the sum. Note that $P < (C-1)N^C$.

Let c_1 have F different prime factors. Then c_2 can be chosen indivisible by some of the factors. Then c_3 can be chosen indivisible by some of the remaining factors and so on. Therefore, we can choose c_i so that $C \leq F + 1$. By [19, V.15.1.b], for the number $\omega(N)$ of distinct prime factors of N, we have $F \leq \omega(N) < 1.39 \ln N / \ln \ln N$. Hence $P < 1.39 \ln N / \ln \ln N \cdot N^{1 + 1.39 \ln N / \ln \ln N}$ and thus $n_0 < N^{4 \ln N - 1}/6$.

$\square$

Both proofs of Lemmata C.5 and C.9 use a technique expressed by the next lemma.

Lemma C.12. *Let $\delta > 0$, $n \in \mathbb{N}$, z_1 be a δ-separated state and $z_1 z_2' \cdots z_n'$ be a feasible path. Then there is a (δ/n)-wide path $z_1 z_2 \cdots z_n$ such that z_n is (δ/n)-separated and for each $1 \leq i \leq n$ we have $z_i \sim z_i'$.*

Proof. For simplicity, we first transform this path into a $\delta/2^n$-wide one. We then show how to improve the result to δ/n-wideness.

For $j \leq n$, we successively construct paths $z_1 \cdots z_j$ that are $\delta/2^j$-wide and z_j is in the same region as z_j' and now is also $(\delta/2^j)$-separated. The state z_1 satisfies all requirements as z_1 is δ-separated. Let $z_1 \cdots z_j$ satisfy the requirements. In particular, $z_j = (s_j, q_j, \nu_j)$ is in the same region as $z_j' = (s_j, q_j, \nu_j')$. Since there is a waiting time t' with $z_j' \xrightarrow{t'} z_{j+1}' = (s_{j+1}, q_{j+1}, \nu_{j+1}')$, there is also an interval of waiting times (a, b) such that for every $t \in (a, b)$ we end up in the same region, i.e. $z_j \xrightarrow{t} z$ for some z of the region containing z_{j+1}'. Moreover, due to $\delta/2^j$-separation of ν_j, we obtain $b - a \geq \delta/2^j$. Therefore, we can choose the waiting time $t = a + \frac{1}{2} \cdot \delta/2^j$ so that also $\nu_j + t$ is $\delta/2^{j+1}$-separated. Hence also $\nu_{j+1} := (\nu_j + t)[\{x \mid \nu_{j+1}'(x) = 0\} := 0]$ is $\delta/2^{j+1}$-separated. We set $z_{j+1} := (s_{j+1}, q_{j+1}, \nu_{j+1})$.

Notice, that this approach guarantees that the fractional parts of the just reset clock are "in the middle" between the surrounding clocks. That is why we needed exponential, i.e. 2^n, deminution of the separation. Nevertheless, due to δ-separation, for every $x, y \in \mathcal{X}$ there are at least n values between $frac(\nu(x))$ and $frac(\nu(y))$ such that even if all were fractional values of other clocks, the state would be δ/n-separated. Also note that as the path is only n steps long, there can be at most n different clocks set between any two clocks. Since we know their ordering in advance, these n different positions are sufficient. $\square$

Now, we can finally start with the promised proofs. Lemma C.5 is a corollary of Lemmata C.11 and C.12.

LEMMA C.5. *Let the region graph $G_{\mathcal{M} \times \mathcal{A}}$ be strongly connected and let p be the period of $G_{\mathcal{M} \times \mathcal{A}}$. Let $k \in \{0, \ldots, p-1\}$, $\delta > 0$ and $R \in V_k$ be a region. Then there is $n \in \mathbb{N}$ such that for every δ-separated $z_1 \in \Gamma_{\mathcal{M} \times \mathcal{A}}^k$ there is a (δ/n)-wide path $z_1 \cdots z_n$ such that z_n is (δ/n)-separated and $z_n \in R$.*
Moreover, we can set $n = \lfloor |V|^{4 \ln |V| - 1}/6 \rfloor \cdot p$.

Proof. In the region graph we have a partition of vertices to sets $V_0, \ldots, V_{p-1}$ due to Lemma 3.12. Let us fix a $k \in \{0, \ldots, p-1\}$. We can define an aperiodic oriented graph (V_k, E_k) where $(R, R') \in E_k$ if there is a path from the region R to the region R' of length exactly p in the region graph $G_{\mathcal{M} \times \mathcal{A}}$.

Let us fix $\delta > 0$ and a region $R \in V_k$. Due to the strong connectedness and aperiodicity of (V_k, E_k) we have by Lemma C.11 in the graph (V_k, E_k) from any region $R' \in V_k$ a path to R of length $x = \lfloor |V|^{4 \ln |V| - 1}/6 \rfloor > \lfloor |V_k|^{4 \ln |V_k| - 1}/6 \rfloor$. Hence, in the graph $G_{\mathcal{M} \times \mathcal{A}}$, we have from R' to R a path of length $n = x \cdot p$.

For every $z_1 \in \Gamma_{\mathcal{M} \times \mathcal{A}}^k$ we have a feasible path $z_1 z_2' \cdots z_n'$ with $z_n' \in R$. We get the (δ/n)-wide path by applying Lemma C.12. $\qquad\square$

LEMMA C.9. *Let the region graph $G_{\mathcal{M} \times \mathcal{A}}$ be strongly connected. Let $\delta > 0$. Let z be a δ-separated state. Then there is $n \in \mathbb{N}$, such that for any $n' \geq n$ there is a state z^* such that the following holds. For any state z_1 almost equal to z there is a (δ/n)-wide path $z_1 \cdots z_{n'}$ such that $z_{n'} = z^*$.*
Moreover, we can set $n = |V| \cdot |\mathcal{X}|$.

Proof. Let $\mathcal{Z}$ be the set of clocks that are not relevant in z. For each clock $x \in \mathcal{Z}$ there is a region R_x such that clock x is reset in region R_x (we make this assumption in Section 3.3). Let us fix a state z_1 almost equal to z. From the strong connectedness we get a feasible path $z_1 z_2' \cdots z_n'$ that for each $x \in \mathcal{Z}$ visits the region R_x. Furthermore, $n \leq |V| \cdot |\mathcal{Z}| \leq |V| \cdot |\mathcal{X}|$. From Lemma C.12 we get a (δ/n)-wide path $z_1 z_2 \cdots z_n$ that also for each $x \in \mathcal{Z}$ visits the region R_x.

Notice that for a fixed state z, control state s and time t there is a unique location q and valuation ν, hence a unique state $z' = (s, q, \nu)$ such that $z \xrightarrow{t} z'$.

Let $t_1, \ldots, t_{n-1}$ be the waiting times and $s_1, \ldots, s_n$ the control states on the path $z_1, \ldots, z_n$. For any $\bar{z}_1$ almost equal to z_1 we can build using the same waiting times and control states a path $\bar{z}_1 \cdots \bar{z}_n$. It is easy to see that for two almost equal states $z, \bar{z}$ a control state s and a time $t > 0$ the states $z', \bar{z}'$ determined by s and t are also almost equal. Inductively, we get that $\bar{z}_i$ is almost equal to z_i for each $1 \leq i \leq n$. Hence, the path $\bar{z}_1 \cdots \bar{z}_n$ is also (δ/n)-wide because δ-wideness is defined only with respect to relevant clocks. We show that $z_n = \bar{z}_n(= z^*)$.

We need a parametrized version of almost equality. For a set of clocks $\mathcal{Y}$ and two states $z = (s, q, \nu)$ and $\bar{z} = (\bar{s}, \bar{q}, \bar{\nu})$ we say that they are $\mathcal{Y}$-equal if $z \sim \bar{z}$ and for each $x \in \mathcal{Y}$ we have $\nu(x) = \bar{\nu}(x)$. The states z_1 and $\bar{z}_1$ are $\mathcal{X}_1$-equal where $\mathcal{X}_1 = \mathcal{X} \setminus \mathcal{Z}$. Let $\mathcal{X}_i$ be a set of clocks and z_i and $\bar{z}_i$ be $\mathcal{X}_i$-equal states. For any $t > 0$ and two states z_{i+1} and $\bar{z}_{i+1}$ such that $z_i \xrightarrow{t} z_{i+1}$ and $\bar{z}_i \xrightarrow{t} \bar{z}_{i+1}$ we have z_{i+1} and $\bar{z}_{i+1}$ are $(\mathcal{X}_i \cup \mathcal{Y})$-equal where $\mathcal{Y}$ is the set of clocks reset in z_i. We get that z_n and $\bar{z}_n$ are $\mathcal{X}$-equal, i.e. $z_n = \bar{z}_n$. It holds because all clocks from $\mathcal{Z}$ are reset on the path $z_1 \cdots z_n$.

Now, for arbitrary $n' \geq n$, we can stretch the path to $z_1 \cdots z_n \cdots z_{n'}$. We get $z_{n'} = \bar{z}_{n'}$ for any starting $\bar{z}_1$ almost equal to z_1 because $z_n = \bar{z}_n$. From same states we can obviously take the same steps to the same successor states. Furthermore, we can easily take (δ/n)-wide transitions by similar arguments as in the proof of Lemma C.4. $\qquad\square$

C.2.4 Proof of Lemma C.7

LEMMA C.7. *Let R be a region. For each $\delta > 0$ there is $\delta' > 0$, $n \in \mathbb{N}$ and $z' \in \Gamma_{\mathcal{M} \times \mathcal{A}}$ such that for every δ-separated $z_1 \in R$ there is a δ'-wide path $z_1 \cdots z_n$ such that z_n and $\bar{z}$ almost equal.*
Moreover, we can set $n = B_{\max} + 1$ and $\delta' = \delta/(B_{\max} + 2)$.

Proof. No relevant clock has in z_1 its fractional value in the interval $(0, \delta)$ because z_1 is δ-separated. We divide this interval into $B_{\max} + 2$ subintervals of equal length and set $\delta' = \delta/(B_{\max} + 2)$.

For a fixed z_1 we inductively build a δ'-wide path $z_1 \cdots z_n$ where $n = B_{\max}+1$. We fix an aribtrary linear order over the set of control states S of the semi-Markov process. Let $1 \leq i < n$. For the state $z_i = (s_i, q_i, \nu_i)$ we choose as s_{i+1} the first state (in the fixed order) such that $\mathbf{P}(s_i)(s_{i+1}) > 0$. This gives us a delay function $f = \mathfrak{D}(s_i, s_{i+1})$. We set b to the integral upper bound of the interval where f is positive if it is not infinity. Otherwise, we set $b = l + 1$ where l is the lower bound of f. Now, we fix the waiting time $t_i = b - \delta'$ and the state $z_{i+1} = (s_{i+1}, q_{i+1}, \nu_{i+1})$ such that $z_i \xrightarrow{t_i} z_{i+1}$.

We show that it is a δ'-wide transition. We divide the set of clocks into two disjunct subsets: the set of clocks $\mathcal{Y}$ that have been reset in one of the states $z_1, \ldots, z_i$ (have been reset at the beginning of the transition to the next state), and all other clocks $\bar{\mathcal{Y}} = \mathcal{X} \setminus \mathcal{Y}$. For each $x \in \mathcal{Y}$ lastly reset in state z_j where $j \leq i$ we have $frac(\nu_{i+1}(x)) = 1 - (i + 1 - j) \cdot \delta'$, i.e. $frac(\nu_{i+1}(x)) \leq 1 - \delta'$ and $frac(\nu_{i+1}(x)) > 1 - \delta > \delta'$. For each $x \in \bar{\mathcal{Y}}$ we have $frac(\nu_{i+1}(x)) = frac(\nu_1(x)) - i \cdot \delta' \geq \delta - i \cdot \delta' \geq \delta'$. Also, $frac(\nu_{i+1}(x)) \leq 1 - \delta - i \cdot \delta' < 1 - \delta'$.

We show that for any δ-separated starting state $\bar{z}_1 \in R$ we reach a state $\bar{z}_n$ almost equal to z_n. We need a parametrized version of almost equality. For a set of clocks $\mathcal{Y}$ and two states $z = (s, q, \nu)$ and $\bar{z} = (\bar{s}, \bar{q}, \bar{\nu})$ we say that they are $\mathcal{Y}$-equal if $z \sim \bar{z}$ and for each $x \in \mathcal{Y}$ we have $\nu(x) = \bar{\nu}(x)$. The states z_1 and $\bar{z}_1$ are $\emptyset$-equal. Let $\mathcal{X}_i$ be a set of clocks and $z_i, \bar{z}_i$ be $\mathcal{X}_i$-equal states. According to the inductive definition, we fix control states $s_{i+1}, \bar{s}_{i+1}$, waiting times $t, \bar{t}$, and states z_{i+1} and $\bar{z}_{i+1}$ such that $z_i \xrightarrow{t} z_{i+1}$ and $\bar{z}_i \xrightarrow{\bar{t}} \bar{z}_{i+1}$. Notice that $s_{i+1} = \bar{s}_{i+1}$, hence $t = \bar{t}$. We have $z_{i+1} \sim \bar{z}_{i+1}$. Furthermore, they are $(\mathcal{X}_i \cup \mathcal{Y})$-equal where $\mathcal{Y}$ is the set of clocks reset in z_i. We get that z_n and $\bar{z}_n$ are almost equal because the paths take at least $B_{\max} + 1 \cdot (1 - \delta') > B_{\max}$ time units. All clocks not reset during this path become irrelevant. We finish the proof by setting $z' = z_n$. $\qquad\square$

D   Proofs of Section 4

D.1   Proof of Theorem 4.1

THEOREM 4.1.   *For every $i \in \mathbb{N}$ we have that*

$$\mathcal{P}_{\mathcal{M}}(\mathcal{R}) - P^i_{\mathcal{M} \times \mathcal{A}}(z_0, \mathcal{C}) \quad \leq \quad \left(1 - \left(\frac{p_{min} \cdot c_{\mathfrak{D}}}{c}\right)^c\right)^{\lfloor i/c \rfloor}$$

where $c = 4 \cdot |V|$.

As $\mathcal{P}_{\mathcal{M}}(\mathcal{R})$ is equal to the probability of reaching $\mathcal{C}$, the i-step transition probabilities $P^i_{\mathcal{M} \times \mathcal{A}}(z, \mathcal{C})$ converge to $\mathcal{P}_{\mathcal{M}}(\mathcal{R})$ as i goes to infinity. Our goal is to show that they converge exponentially quickly.

Our proof proceeds as follows. Denote by B the union of all regions that belong to BSCCs of $G_{\mathcal{M} \times \mathcal{A}}$. We show that for $c = 4 \cdot |V|$ there is a lower bound $p_{bound} > 0$ on the probability of reaching B in at most c steps from any state $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}$. Note that then the probability of not hitting B after $i = m \cdot c$ steps is at most $(1 - p_{bound})^m$. However, this means that $P^i_{\mathcal{M} \times \mathcal{A}}(z, \mathcal{C})$ cannot differ from the probability of reaching $\mathcal{C}$ (and thus also from $\mathcal{P}_{\mathcal{M}}(\mathcal{R})$) by more than $(1 - p_{bound})^m$ because $\mathcal{C} \subseteq B$ and the probability of reaching $\mathcal{C}$ from $B \smallsetminus \mathcal{C}$ is 0. Moreover, we show that p_{bound} can be set to $(p_{\min} \cdot c_{\mathfrak{D}} \cdot 1/c)^c$, from which we obtain the desired upper bound on $|\mathcal{P}_{\mathcal{M}}(\mathcal{R}) - P^i_{\mathcal{M} \times \mathcal{A}}(z, A)|$.

So to obtain the desired result, it suffices to prove the following

Proposition D.1. *For every $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}$ we have that*

$$P^c_{\mathcal{M} \times \mathcal{A}}(z, B) \quad \geq \quad p_{bound}$$

Here $c = 4 \cdot |V|$ and $p_{bound} = (p_{\min} \cdot c_{\mathfrak{D}} \cdot 1/c)^c$.

Note that this section draws heavily on some of the methods and lemmas proved in the previous section, though often in a slightly easier form. However, to keep individual sections of the Appendix independent, we repeat the arguments here once more.

Similarly to previous section, we are interested in paths $z \ldots z_n$ that are δ-wide. For a fixed $\delta > 0$, it means that the waiting time of any transition in the path can be changed by $\pm \delta$ without ending up in a different region in the end. Precise definition follows.

Definition D.2. *Let $z = (s, q, \nu)$ and $z' = (s', q', \nu')$ be two states. For a waiting time $t \in \mathbb{R}_{>0}$ we set $z \xrightarrow{t} z'$ if $\mathcal{A}(z) = (q', \bar{\nu})$ and $\nu' = \bar{\nu} + t$. We set $z \to z'$, called a feasible transition, if for some $t \in \mathbb{R}_{>0}$ (i) $z \xrightarrow{t} z'$; and (ii) $f_d(t) > 0$, where $f_d = \mathcal{D}(s, s')$.*

For $\delta > 0$, we say that a feasible transition $z \to z'$ is δ-wide if for every $x \in \mathcal{X}$ relevant for ν' we have $frac(\nu_i(x)) \in [\delta, 1 - \delta]$.

Let $z_1 \cdots z_n$ be a path. It is feasible if for each $1 \leq i < n$ we have that $z_i \to z_{i+1}$. It is δ-wide if for each $1 \leq i < n$ we have that $z_i \to z_{i+1}$ is a δ-wide transition.

We first show that any δ-wide path of a finite length, say n, from any state $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}$ to a state z_n in a region R, induces a set of paths from z to the region R, and that their probability is bounded below by a positive constant.

Lemma D.3. *For every $\delta > 0$ and $n > 1$ there is $b > 0$ such that the following holds. For every δ-wide path $\sigma = z_0 z_1 \cdots z_n$, there is a set of states $Z \ni z_n$ such that it holds $P^n_{\mathcal{M} \times \mathcal{A}}(z_1, Z) \geq b$.*
Moreover, we can set $b = (p_{\min} \cdot c_{\mathfrak{D}} \cdot 2\delta/n)^n$.

Proof. We fix any δ-wide path $\sigma = z_0 z_1 \cdots z_n = (s_0, q_0, \nu_0)(s_1, q_1, \nu_1) \cdots (s_n, q_n, \nu_n)$. For $1 \leq i \leq n$, let t_i be the waiting times such that $z_{i-1} \xrightarrow{t_i} z_i$, and let $X_i = \{x \in \mathcal{X} \mid \mathcal{A}(z_{i-1}) = (q, \nu), \ \nu(x) = 0\}$ be the set of clocks reset right before waiting t_i.

For $\varepsilon > 0$, we define an ε-neighbourhood of σ to be the set of paths of the form $z_0 \xrightarrow{t_1'} (s_1, q_1, \nu_1') \cdots \xrightarrow{t_n'} (s_n, q_n, \nu_n')$ where $t_i' \in (t_i - \varepsilon, t + \varepsilon)$. Due to δ-wideness of σ, all paths of its δ/n-neighbourhood are feasible, and follow the same sequence of regions. Considering this δ/n-neighbourhood, the set of all possible ν_n's forms the sought set of states Z.

We now give a lower bound on $P^n_{\mathcal{M} \times \mathcal{A}}(z, Z)$. First, recall the following notation: let $p_{\min}$ denote the smallest probability in $\mathcal{M}$. Further, let us denote by $\mathfrak{D}(\mathcal{M})$ the set of delay densities used in $\mathcal{M}$, i.e. $\mathfrak{D}(\mathcal{M}) = \{\mathbf{D}(s, s') \mid s, s' \in S\}$. From our assumptions imposed on delay densities we obtain the following uniform bound $c_{\mathfrak{D}} > 0$ on delay densities of $\mathfrak{D}(\mathcal{M})$. For every $f \in \mathfrak{D}(\mathcal{M})$ and for all $x \in [0, B_{\max}]$, either $f(x) > c_{\mathfrak{D}}$ or $f(x) = 0$, and moreover, $\int_{B_{\max}}^{\infty} f(x) dx > c$ or equals 0.

We define sets of states $Z_0, Z_1 \ldots, Z_n = Z$, where Z_i is the set of all states (s_i, q_i, ν_i') in the δ-neighbourhood of σ. Note that $P_{\mathcal{M} \times \mathcal{A}}(z_0, Z_1) = \mathbf{P}(s_0)(s_1) \cdot \int_{t_0 - \delta/n}^{t_0 + \delta/n} f_d(t) dt$, where f_d is the appropriate delay density for this transition. Using the bounds given above, $P_{\mathcal{M} \times \mathcal{A}}(z_0, Z_1) \geq p_{\min} \cdot \int_{t_0 - \delta/n}^{t_0 + \delta/n} c_{\mathfrak{D}} dt = p_{\min} \cdot c_{\mathfrak{D}} \cdot 2\delta/n$. Similarly, for any $z_i' \in Z_i$, $P_{\mathcal{M} \times \mathcal{A}}(z_i', Z_{i+1}) \geq p_{\min} \cdot c_{\mathfrak{D}} \cdot 2\delta/n$ holds by the same arguments. Therefore, from the definition of the n-step transition kernel, $P^n_{\mathcal{M} \times \mathcal{A}}(z_0, Z) \geq (p_{\min} \cdot c_{\mathfrak{D}} \cdot 2\delta/n)^n$. $\qquad\square$

We now prove that from any state $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}$, some BSCC reachable from z in the region graph is also reachable from z along a δ-wide path, and that this path length is bounded from above by a constant.

We use two steps: first, we show that, from any $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}$, we can reach a δ-*separated* state along δ'-wide path of bounded length; second, once in a δ-separated state, we construct a δ''-wide path of length at most $|V|$ ending in the BSCC.

Definition D.4. *Let $\delta > 0$. We say that a set $X \subseteq \mathbb{R}_{\geq 0}$ is δ-separated if for every $x, y \in X$ either $frac(x) = frac(y)$ or $|frac(x) - frac(y)| > \delta$.*
Further, we say that $(s, q, \nu) \in \Gamma_{\mathcal{M} \times \mathcal{A}}$ is δ-separated if the set

$$\{0\} \cup \{\nu(x) \mid x \in \mathcal{X}, x \text{ is relevant for } \nu\}$$

is δ-separated.

Lemma D.5. *There is $\delta > 0$ and $n \in \mathbb{N}$ such that for any $z_1 \in \Gamma_{\mathcal{M} \times \mathcal{A}}$ there is a δ-wide path $z_1 \cdots z_n$ such that z_n is δ-separated.*
Moreover, we can set $n = B_{\max} \cdot (|\mathcal{X}| + 2)$ and $\delta = 1/(2(|\mathcal{X}| + 2))$.

Proof. (Same as Lemma C.4) To simplify the argumentation we introduce a notion of a *r-grid* that marks r distinguished points (called *lines*) on the $[0, 1]$ line segment. In the proof we show that we can place fractional values of all relevant clocks on such distinguished points. Let $r \in \mathbb{N}$. We say that a set of clocks $\mathcal{Y} \subseteq \mathcal{X}$ is *on r-grid* in z if for every $x \in \mathcal{Y}$ relevant in z we have $frac(\nu(x)) = n/r$ for some $0 \le n < r$. For $0 \le n < r$, we say that the *n-th line of the r-grid is free* in z if there is no relevant clock in the $1/2k$-neighborhood of the n-th line, i.e. for any relevant $x \in \mathcal{X}$ we have $frac(\nu(x)) \notin (n/r - 1/2r, n/r + 1/2r)$.

Let $r = |X| + 2$. We inductively build a $1/2r$-wide path $z_1 \cdots z_n$ where $n = B_{\max} \cdot r$. The set $\emptyset$ is on r-grid in z_1. We show that if a set $\mathcal{Y}_i$ is on r-grid in state z_i, there is a $1/2k$-wide transition to z_{i+1} such that $(\mathcal{Y}_i \cup \mathcal{Z})$ is on r-grid in z_{i+1} where $\mathcal{Z}$ is the set of clocks newly reset in z_i. There are $|X| + 2$ lines on the grid and only $|X|$ clocks. At least two of these lines must be free. Let $j \ne 0$ be such a line. Let t be a waiting time and z_{i+1} a state such that $frac(t) = 1 - j/r$ and $z_i \xrightarrow{t} z_{i+1}$. Such waiting time must be indeed possible because the interval where the density function of any transition is positive has integral bounds. The transition $z_i \xrightarrow{t} z_{i+1}$ is $1/2r$-wide because the line j is free in z_i. Furthermore, the set $(\mathcal{Y}_i \cup \mathcal{Z})$ is on r-grid in z_{i+1} because the fractional value of each clock that was previously on r-grid was changed by a multiple of $1/r$. The newly reset clocks have fractional value $1 - j/r$ which is again a multiple of $1/r$.

Next, we show that $\mathcal{X}$ is on r-grid in z_n. Clocks reset in this path on r-grid in z_n. The remaining clocks are all irrelevant because the path of $B_{\max} \cdot r$ steps takes at least $B_{\max}$ time units. Indeed, each transition in this path takes at least $1/r$ time unit. According to the definition, $\mathcal{X}$ is on r-grid in z_n. Hence, the state z_n is $1/r$-separated because the distance between two adjacent grid lines is $1/r$. By setting $\delta = 1/2r$ we get the result. $\square$

Lemma D.6. *Let $\delta, \delta' > 0$ and R be a region. Then there is $n \in \mathbb{N}$ such that for every δ-separated $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}$ it holds that if there is a feasible path from z to z', for a z' in the region R, then there is also $i \le n$ and a δ'-wide path $z \cdots z_i$ such that $z_i \in \Gamma_{\mathcal{M} \times \mathcal{A}} \cap R$ is δ'-separated.*
Moreover, we can set $n = |V|$ and $\delta' = \delta/|V|$.

Proof. For simplicity, we first transform this path into a $\delta/2^n$-wide one. We then show how to improve the result to δ/n-wideness.

Let us fix any δ-separated state $z \in \Gamma_{\mathcal{M} \times \mathcal{A}}$, belonging to a particular region, say R_s. We will show that for any region R_x such that $(R_x, R_s) \in E$ in the region graph, we can find a waiting time t and δ-separated state z_1 belonging to R_x, such that $z \xrightarrow{t} z_1$.

As R_x is reachable from R_s in one step in the region graph, there is an interval of waiting times (a, b) such that for every $t' \in (a, b)$ $z \xrightarrow{t} z'_1$ for some z'_1 from R_x. Moreover, due to δ-separation of z, we obtain $b - a \ge \delta$. Therefore, we can choose the waiting time $t = (a + b)/2$ and z'_1 is $\delta/2$-separated. Intuitively, we need to 'lower' the δ-separation and wideness in each step as we might be

forced to reset a clock, say x_r, to a place between two other clocks, say x_1, x_2, with $|frac(x_1 - x_2)| = \delta$.

Note that if the state z' is reachable from z along a feasible path, it must be also reachable in at most $|V|$ steps in the region graph. In such case, we can put $n = |V|$ and the δ' would be equal to $\delta/2^n$. However, due to δ-separation, for every $x, y \in \mathcal{X}$ there are at least n values between $frac(\nu(x))$ and $frac(\nu(y))$ such that even if all were fractional values of other clocks, the state would be δ/n-separated. Also note that as the path is only n steps long, there can be at most n different clocks set between any two clocks. Since we know their ordering in advance, these n different positions are sufficient, and we can set $\delta' = \delta/|V|$. □

Now we are ready to prove the Proposition D.1.

of Proposition D.1. Lemma D.5 together with Lemma D.6 give us an upper bound on the number of steps c_b needed to hit a state in one of the BSCCs along a δ-wide path from any state in $\Gamma_{\mathcal{M} \times \mathcal{A}}$: we can set $c_b = B_{\max} \cdot (|\mathcal{X}| + 2) + |V|$ and $\delta = (1/(2 \cdot (|\mathcal{X}| + 2)))$. From Lemma D.3 we have

$$P_{\mathcal{M} \times \mathcal{A}}^{c_b}(z, B) \quad \geq \quad \left(\frac{p_{\min} \cdot c_{\mathfrak{D}}}{2(|\mathcal{X}| + 2) \cdot c_b} \right)^{c_b}$$

As $c_b \leq 2 \cdot |V|$ for all but very small region graphs we have

$$\geq \quad \left(\frac{p_{\min} \cdot c_{\mathfrak{D}}}{2(|\mathcal{X}| + 2) \cdot 2 \cdot |V|} \right)^{2 \cdot |V|}$$

$$\geq \quad \left(\frac{p_{\min} \cdot c_{\mathfrak{D}}}{(4 \cdot |V|)^2} \right)^{2 \cdot |V|}$$

$$\geq \quad \left(\frac{p_{\min} \cdot c_{\mathfrak{D}}}{4 \cdot |V|} \right)^{4 \cdot |V|}$$

From this, we get the desired

$$P_{\mathcal{M} \times \mathcal{A}}^{c}(z, B) \quad \geq \quad \left(\frac{p_{\min} \cdot c_{\mathfrak{D}}}{c} \right)^{c}$$

where $c = 4 \cdot |V|$. □